

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 8, 2001, 06:21:22 ; Search time 217.33 seconds
(without alignments)
53.723 Million cell updates/sec

Title: US-09-507-242-3

Perfect score: 20
Sequence: 1 ccagatgataaagtagaac 20

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 678276 seqs, 291890651 residues
Total number of hits satisfying chosen parameters: 1356552

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0401:*

1: /SID56/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseqn/NA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	21	AS0293
2	20	100.0	989	21	AS0292
3	16.4	82.0	19	21	AS3104
4	15.8	79.0	2224	18	V74796
5	15.4	77.0	19	21	AS3103
6	15.2	76.0	470	18	T75261
7	15.2	76.0	981	21	C44949
8	15.2	76.0	1434	21	C45997
9	15.2	76.0	1606	13	Q29162
10	15.2	76.0	1792	21	C46375
11	15.2	76.0	1797	21	C34601

C	12	15.2	76.0	1892	21	C52084
C	13	15.2	76.0	2329	18	T61274
C	14	15.2	76.0	4780	16	O94253
C	15	15	75.0	653	20	V87826
C	16	14.8	74.0	42	11	O04070
C	17	14.8	74.0	42	14	O42527
C	18	14.8	74.0	312	19	V30233
C	19	14.8	74.0	792	21	C54232
C	20	14.8	74.0	923	21	C35559
C	21	14.8	74.0	1563	21	A99072
C	22	14.8	74.0	2058	21	C46199
C	23	14.8	74.0	2058	16	O75303
C	24	14.8	74.0	4247	20	X00915
C	25	14.8	74.0	5156	21	Z98324
C	26	14.8	74.0	6828	16	O75302
C	27	14.4	72.0	1132	21	C36483
C	28	14.4	72.0	1626	21	C74280
C	29	14.4	72.0	2220	21	C79691
C	30	14.4	72.0	65921	21	Z89046
C	31	14.4	72.0	58073	18	T58840
C	32	14.2	71.0	227	21	A44514
C	33	14.2	71.0	251	19	X11239
C	34	14.2	71.0	251	19	X11240
C	35	14.2	71.0	260	21	C43927
C	36	14.2	71.0	276	19	V26672
C	37	14.2	71.0	298	21	A01212
C	38	14.2	71.0	304	21	A43467
C	39	14.2	71.0	304	21	C98398
C	40	14.2	71.0	363	21	C08176
C	41	14.2	71.0	514	21	C37256
C	42	14.2	71.0	585	21	C52247
C	43	14.2	71.0	600	20	X61550
C	44	14.2	71.0	639	22	C84300
C	45	14.2	71.0	651	20	X61557

ALIGNMENTS

RESULT 1	
AS0293	standard; DNA; 20 BP.
AC	AS0293;
XX	
DT	20-NOV-2000 (first entry)
XX	
DE	Candida albicans CaESS1 gene-specific primer OW-216.
XX	
KW	CaESS1; ESS1; infection; diagnosis; therapy; antifungal; fungicide;
KW	antiproliferative; cytostatic; antitumor; PCR primer; ss.
XX	
OS	Candida albicans.
XX	
PN	WO2000050561-A2.
XX	
PD	31-AUG-2000.
XX	
PF	18-FEB-2000; 2000WO-US04203.
XX	
PR	23-FEB-1999; 99US-0121246.
XX	
PA	(HEAL-) HEALTH RES INC.
XX	
PI	Hanes SD, Devasahayam G, Chaturvedi V;
XX	
DR	WPI: 2000-565453/52.
XX	
PT	Novel Candida albicans gene, CaESS1 useful for identifying compounds
PT	that specifically bind to and/or inhibit CaESS1 and thus for treating
PT	Candida albicans infections and other life-threatening fungal
PT	infections
XX	

Arabidopsis thalia
Penicillium decumb
Neuronal nitrogen
EST clone EM13 H
Probe to endothel
DNA probe for HET3
cDNA encoding per1
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
Human astrovirus s
protease activated
A. thaliana gene 1
Genomic RNA of hum
Arabidopsis thalia
Human secreted pro
Human secreted pro
Human nitrin DNA
Mycoplasma genital
Human secreted exp
Human diallelic po
Human diallelic po
Arabidopsis thalia
Arabidopsis thalia
B. burgdorferi ant
Human EXCS encodin
B. burgdorferi ant

PS Claim 6; Page 38; 51pp; English.

CC This is the nucleotide sequence of PCR primer OW-216, which is
CC specific to the CaESS1 gene (see A50292), the Candida albicans
CC homologue of Saccharomyces cerevisiae ESS1. OW-216 can be used
CC with primer OW-221 (see A50294) for the specific amplification of
CC CaESS1 DNA; they do not amplify ESS1 DNA or PIN1 (the human
CC homologue of ESS1) DNA. Detection of CaESS1 nucleic acids can
CC be used as a means for diagnosing C. albicans infection.

XX Sequence 20 BP; 8 A; 3 C; 5 G; 4 T; 0 other;

QY Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ccagatggtataagtagaac 20
|||||

RESULT 2

A50292
ID A50292 standard; DNA; 989 BP.
XX
AC A50292;
XX
DT 20-NOV-2000 (first entry)
XX
DE Candida albicans CaESS1 gene.
XX
KW CaESS1; ESS1; infection; diagnosis; therapy; antifungal; fungicide;
KW antiproliferative; cytostatic; antitumour; ds.
XX
OS Candida albicans.
XX
FH Key Location/Qualifiers
FT CDS 256..789
FT /*tag= a

XX WO2000050561-A2.
XX
PD 31-AUG-2000.
XX
PF 18-FEB-2000; 2000WO-US04203.
XX
PR 23-FEB-1999; 99US-0121246.
XX
PA (HEAL-) HEALTH RES INC.
XX
PI Hanes SD, Devasahayam G, Chaturvedi V;
XX
DR WPI: 2000-565453/52.
DR P-PSDB; Y95876.
XX
XX Novel Candida albicans gene, CaESS1 useful for identifying compounds
PT that specifically bind to and/or inhibit CaESS1 and thus for treating
PT Candida albicans infections and other life-threatening fungal
PT infections -
XX
XX Claim 2; Fig 1A; 51pp; English.

PS This is the complete nucleotide sequence of the CaESS1 gene from
XX Candida albicans. The gene encodes a 177-amino acid protein,
CC CaESS1 (see Y95876), which is the C. albicans homologue of
CC Saccharomyces cerevisiae ESS1 (428 amino acid identity). The
CC CaESS1 gene was isolated from a C. albicans genomic DNA library by
CC functional complementation of a temperature-sensitive S. cerevisiae
CC strain, ess1-194pts. CaESS1 nucleic acids, especially CaESS1-specific
CC primers (see A50293-94) and probes to determine the presence of C.
CC albicans in a sample or specimen. CaESS1 protein is a target for
CC screening for antifungal and inhibitor compounds, useful for

CC treating or preventing C. albicans infections. CaESS1 DNA can also
CC be used to generate diagnostic probes or primers for replicating or
CC cloning C. albicans DNA. A CaESS1 inhibitor can be a compound
CC which selectively inhibits growth of S. cerevisiae not containing an
CC endogenous ESS1 gene but rather CaESS1 and uninduced PIN1 (a human
CC homologue of ESS1) and/or preferably does not inhibit induced PIN1,
CC e.g. does not inhibit S. cerevisiae not containing an endogenous
CC ESS1 but rather induced PIN1. Compositions which inhibit PIN1
CC are useful antiproliferatives e.g. antineoplastics, antitumour
CC agents or anticancer agents.

XX Sequence 989 BP; 340 A; 174 C; 190 G; 285 T; 0 other;

QY Query Match 100.0%; Score 20; DB 21; Length 989;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 ccagatggtataagtagaac 20
|||||

Db 520 ccagatggtataagtagaac 539

RESULT 3

A83104
ID A83104 standard; DNA; 19 BP.
XX
AC A83104;
XX
DT 04-DEC-2000 (first entry)
XX
DE cdk7 ribozyme binding site #25.
XX
KW Ribozyme; hairpin; hammerhead; gene therapy; vasotropic;
KW restenosis; ss.
XX
OS Mammalia.
XX
PN WO200032765-A2.
XX
PD 08-JUN-2000.
XX
PF 06-DEC-1999; 99WO-US28772.
XX
PR 04-DEC-1998; 98US-0110954.
XX
PA (IMMU-) IMMUSOL INC.
XX
PI Tritz R, Welch PJ, Barber JR, Robbins JM;
XX
DR WPI: 2000-412314/35.
XX
PT New hairpin and hammerhead ribozyme for inhibiting restenosis, cleaves
PT RNA encoding a cyclin or cell-cycle dependent kinase other than CDK1,
PT PCNA and Cyclin B1 -
XX
XX Disclosure; Page 56; 109pp; English.

PS The present invention relates to a hairpin or hammerhead ribozyme,
XX designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC Representative examples of ribozyme recognition sites are given in
CC A82415 to A86787. The ribozyme of the invention is useful for
CC inhibiting restenosis by introduction of the ribozyme into cells.
CC The ribozyme is resistant to endonuclease activity and hence is
CC efficient in restenosis treatment.

XX Sequence 19 BP; 9 A; 2 C; 4 G; 4 T; 0 other;

QY Query Match 82.0%; Score 16.4; DB 21; Length 19;
Best Local Similarity 94.4%; Pred. No. 24;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	3	agatggtataagtagaac	20
Db	1	agatggtataaataagaac	18

RESULT	4
V74796/c	
ID	V74796 standard; DNA; 2224 BP.

DT 16-MAR-1999 (first entry)

AA	DE	Staphylococcus aureus contig	SEQ ID #485

XX Computer readable medium; vaccine; S.aureus infection; immunodetection
 KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
 KW skin infection; surgical wound infection; scalded skin syndrome;
 KW toxic shock syndrome; ds.

Staphylococcus aureus.

	Location/Qualifiers
FH	Key
FT	misc_feature
	1681..1740

FT	/"ag	"these bases represent a line of missing text in
FT	/note-	the sequence listing in the specification. They
FT		are included to maintain the nucleotide numbering
FT		given in the specification for this DNA sequence"

PN EP786519-A2.

PD 30-JUL-1997

PF 07-JAN-1997; 97EP-0100117.

PR 05-JAN-1996; 9605-0009861

PA (HUMA-) HUMAN GENOME SCL INC.

PI BARASH SC, CHOI GH, DILLON FO, LAMMON AM, ...
 PI BOGOS CA: ...

DR WPI: 199/-3/4922/35

aa polynucleotide(s) and proteins derived from *Staphylococcus aureus*
 PT stored on computer readable medium and used in the production of
 PT anti-*S. aureus* vaccines

PS Claim 1; Page 1412-1413; 327Lpp; English...

xx This sequence represents one of 5191 *Staphylococcus aureus* DNA sequences of the invention. The DNA sequences are recorded on a computer readable medium, preferably selected from a floppy or hard disk, random access memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using the *S. aureus* DNA sequences allows putative functions to be assigned to that protein-encoding or regulatory regions of commercial, therapeutic or industrial importance can be obtained. Specifically, sequences which are likely to encode antigens have been identified and these polypeptides can be used in a vaccine composition against *S. aureus* infection. The polypeptides can also be used in a kit for the immunodetection of *S. aureus* in a sample. *S. aureus* is implicated in numerous human diseases, including cellulitis, eyelid infections, food poisoning, osteomyelitis, skin and surgical wound infections, scalded skin syndrome, toxic shock syndrome, etc. Organisms transformed with the DNA sequences can be used for recombinant production of the polypeptides. The new DNA sequences (and their fragments) are useful as primers or probes for isolating homologues of any of the *S. aureus* DNA sequences contained on the computer readable medium.

Sequence 2224 BP; 785 A; 291 C; 435 G; 648 T; 65 other;

Query Match	79.0%	Score 15.8;	DB 18;	length 2224;
Best Local Similarity	89.5%;	Pred. No. 70;		
Matches 17;	Conservative	0;	Mismatches 2;	Indels 0;
				Gaps 0;

QY	2	cagatggtataagtagaac	20
Db	796	CAGATGCTTAAGTACAAC	778

RESULT	5
A83103	
ID	A83103 standard; DNA; 19 BP.

AC A83103;

DT 04-DEC-2000 (first entry)

DE cdk7 ribozyme binding site #24.

Ribozyme; hairpin; hammerhead; gene therapy, vasculoprotection

XX

XX
XX
W0300032765-A2

XX
09-TIN-2000XX
XX
06-DEC-1999

XX
PR 04-DEC-1998: 98US-0110954.

XX
PA
(IMMU-) IMMUSOL INC.

AA Tritz R, Welch PJ, Barber JR, Robbins JM;
PI

WPI; 2000-412314/35.

PT New hairpin and hamr

PT PCNA and Cyclin B1

PS Disclosure; Page 56; 109pp; English.

AA The present invention relates to a hairpin orhammerhead ribozyme,
CC designed to cleave RNA encoding a cyclin or cell-cycle dependent kinase
CC other than cell-cycle dependent kinases CDK1, PCNA and Cyclin B1.
CC Representative examples of ribozyme recognition sites are given in
CC A8415 to A86787. The ribozyme of the invention is useful for
CC inhibiting restenosis by introduction of the ribozyme into cells.
CC The ribozyme is resistant to endonuclease activity and hence is
CC efficient in restenosis treatment.

Sequence 19 BP; 11 A; 0 C; 4 G; 4 T; 0 other

Query Match	77.03;	Score 15.4;	DB 21;	Length 19;
Best Local Similarity	94.18;	Pred. No. 76;		
Matches 16; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	3	agatggtataagtagaa	19
Db	3	agatggtataaataagaa	19

RESULT	6
T75261/c	
ID T75261	standard; DNA; 470 BP.

AC T75261

DT 21-DEC-1998 (first entry)

XX Nucleotide sequence encoding human RAD50 Intron 10.
 DE ds; human; RAD50; DNA repair; tumour suppression; cancer; Septin-2;
 XX central nervous system.
 XX Homo sapiens.
 OS M09727284-A2.
 PN 31-JUL-1997.
 PD 24-JAN-1997; 97WO-US01299.
 XX 17-JUL-1996; 96US-0687080.
 PR 26-JAN-1996; 96US-0592126.
 XX (GENE-) GENELABS TECHNOLOGIES INC.
 PA Dolganov G;
 PI WPI; 1997-393672/36.
 DR Human tumour suppressor gene RAD50 - useful to detect
 XX predisposition to, decrease risk of and treat cancer, also Septin-2
 PT homologues
 XX Claim 1; Page 111-112; 195pp; English.
 XX The human RAD50 (hrRAD50) is involved in DNA repair and has tumour
 CC suppression activity, can be used to detect predisposition to, decrease
 CC the risk of or treat cancers, e.g. acute myeloid leukaemia,
 CC myelodysplastic syndrome, therapy related myelodysplastic syndrome,
 CC therapy related acute myeloid leukaemia, refractory anaemia or refractory
 CC anaemia with excess blasts. Also disclosed in this invention is Human
 CC Septin-2 homologues of which may be used as targets for cancer therapies
 CC and central nervous system directed treatment methods, and to measure the
 CC proliferative potential of selected cell types.
 CC
 XX Sequence 470 BP; 154 A; 60 C; 60 G; 196 T; 0 other;
 SQ

Query Match 76.0%; Score 15.2; DB 18; Length 470;
 Best Local Similarity 85.0%; Pred. NO. 1.2e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccaagtggtataagtagaac 20
 Db 37 CCAGCAGCATATATTAGAAC 18
 ||||| ||||| |||||

RESULT 7
 C44949/c
 ID C44949 standard; DNA; 981 BP.
 XX C44949;
 AC
 XX 18-OCT-2000 (first entry)
 DT
 XX Arabidopsis thaliana DNA fragment SEQ ID NO: 44735.
 DE
 XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX Arabidopsis thaliana.
 OS
 PN EPI033405-A2.
 PD 06-SEP-2000.
 PF 25-FEB-2000; 2000EP-0301439.
 XX

PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126284.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139753.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158332.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 981;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ccagatggtataagtagaac 20
||||| ||| ||||| |||||
Db 650 CCAGAGGTTTAACTTGAAAC 631

RESULT 8
C45997/c
ID C45997 standard; DNA; 1434 BP.
XX

AC C45997;
XX

DT 18-OCT-2000 (first entry)
XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48536.
XX

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.
XX

PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX

PF 25-FEB-2000; 2000EP-0301439.
XX

XX 25-FEB-1999; 99US-01231825.
PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140635.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144864.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148664.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

```

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159368.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          76.0%; Score 15.2; DB 21; Length 1434;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy 1 ccagatggtataagtagaac 20
    ||||| ||||| |||||
Db 1103 CCAGAGCTTTAAGTTGAGAC 1084

```

```

RESULT 9
Q29162/c
ID Q29162 standard; DNA; 1606 BP.

```

```

AC Q29162;
DT 05-MAR-1993 (first entry)

```

```

DE Human R3BP.

```

```

KW Gastrin releasing peptide; GRP; bombesin; neuromedin B; NMB;
KW ranatensin; RBP; bombesin-like peptide; R1BP; R2BP; R3BP;
KW receptor; agonist; antagonist; ligand; antibody; cancer; ss.

```

```

OS Homo sapiens.

```

```

FH Key Location/Qualifiers
FT CDS 172..1371
    /*tag= a

```

```

PN W09216623-A.

```

```

PD 01-OCT-1992.

```

```

PF 13-MAR-1992; 92WO-US02091.

```

```

PR 15-MAR-1991; 91US-0670603.

```

```

PR 03-OCT-1991; 91US-0771332.

```

```

PA (BERL.) BERLEX LAB INC.

```

```

PA (USSH) US DEPT HEALTH & HUMAN SERVICE.

```

```

PI Battey JF, Corlay MH, Fathi Z, Feldman RI, Harkin RN;

```

```

PI Slattey TK, Wada E, Wu JM;
XX WPI; 1992-349208/42.

```

```

DR P-PSDB; R27511.
XX
XX Receptors for bombesin-like peptide(s) and their DNA - useful for
PT screening for agonists and antagonists of the receptor ligands,
PT also for treating or diagnosing cancer
XX
XX Disclosure: Page 147-149; 173pp; English.

```

```

XX The DNA sequences encoding mouse R1BP; human R1BP; rat R2BP; human
CC R2BP and human R3BP are given in Q29158-62 respectively. The
CC receptor gene and encoded polypeptide are used for screening for
CC agonists and antagonists of the receptor ligands, for producing
CC diagnostic or therapeutic reagents, and for producing antibodies.
CC Hosts suffering from abnormal receptor function, e.g. proliferative
CC cell conditions such as cancers, may be treated.
CC The mouse GRP receptor was isolated from Swiss 3T3 fibroblasts and
CC sequenced. The sequence was used to design oligonucleotide probes
CC to isolate DNA encoding mouse GRP receptor from a Swiss 3T3 CDNA
CC library. This DNA was then used as a probe to isolate rat NMB
CC receptor, human GRP receptor, human NMB receptor and human R3BP
CC (incompletely characterised homologous putative receptor) from
CC DNA libraries.

```

```

SQ Sequence 1606 BP; 428 A; 367 C; 323 G; 488 T; 0 other;

```

```

Query Match          76.0%; Score 15.2; DB 13; Length 1606;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

Oy 1 ccagatggtataagtagaac 20
    ||||| ||||| |||||
Db 87 CGAGATGCTTATGAGAGAC 68

```

```

RESULT 10
C46375/c
ID C46375 standard; DNA; 1792 BP.

```

```

AC C46375;

```

```

DT 18-OCT-2000 (first entry)

```

```

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49913.

```

```

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 49913.

```

```

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

```

```

OS Arabidopsis thaliana.

```

```

XX EP1033405-A2.

```

```

XX 06-SEP-2000.

```

```

PF 25-FEB-2000; 2000EP-0301439.

```

```

PR 25-FEB-1999; 99US-0121825.

```

```

PR 05-MAR-1999; 99US-0123180.

```

```

PR 09-MAR-1999; 99US-0123548.

```

```

PR 23-MAR-1999; 99US-0125788.

```

```

PR 25-MAR-1999; 99US-0126264.

```

```

PR 29-MAR-1999; 99US-0126785.

```

```

PR 01-APR-1999; 99US-0127462.

```

```

PR 06-APR-1999; 99US-0128234.

```

```

PR 08-APR-1999; 99US-0128714.

```

```

PR 16-APR-1999; 99US-0129845.

```

```

PR 19-APR-1999; 99US-0130077.

```

```

PR 21-APR-1999; 99US-0130449.

```

```

PR 23-APR-1999; 99US-0130510.

```

```

PR 28-APR-1999; 99US-0130891.

```

```

PR 30-APR-1999; 99US-0132048.

```

PR 30-APR-1999; 9905-0132407.
 PR 04-MAY-1999; 9905-0132484.
 PR 05-MAY-1999; 9905-0132485.
 PR 06-MAY-1999; 9905-0132486.
 PR 06-MAY-1999; 9905-0132487.
 PR 07-MAY-1999; 9905-0132863.
 PR 11-MAY-1999; 9905-0134256.
 PR 14-MAY-1999; 9905-0134218.
 PR 14-MAY-1999; 9905-0134219.
 PR 14-MAY-1999; 9905-0134370.
 PR 18-MAY-1999; 9905-0134768.
 PR 19-MAY-1999; 9905-0134941.
 PR 20-MAY-1999; 9905-0135124.
 PR 21-MAY-1999; 9905-0135353.
 PR 24-MAY-1999; 9905-0135629.
 PR 25-MAY-1999; 9905-0136021.
 PR 27-MAY-1999; 9905-0136392.
 PR 28-MAY-1999; 9905-0136782.
 PR 01-JUN-1999; 9905-0137222.
 PR 03-JUN-1999; 9905-0137528.
 PR 04-JUN-1999; 9905-0137502.
 PR 07-JUN-1999; 9905-0137724.
 PR 08-JUN-1999; 9905-0138094.
 PR 10-JUN-1999; 9905-0138540.
 PR 10-JUN-1999; 9905-0138847.
 PR 14-JUN-1999; 9905-0139119.
 PR 16-JUN-1999; 9905-0139452.
 PR 16-JUN-1999; 9905-0139453.
 PR 17-JUN-1999; 9905-0139492.
 PR 18-JUN-1999; 9905-0139454.
 PR 18-JUN-1999; 9905-0139455.
 PR 18-JUN-1999; 9905-0139456.
 PR 18-JUN-1999; 9905-0139457.
 PR 18-JUN-1999; 9905-0139458.
 PR 18-JUN-1999; 9905-0139459.
 PR 18-JUN-1999; 9905-0139460.
 PR 18-JUN-1999; 9905-0139461.
 PR 18-JUN-1999; 9905-0139462.
 PR 18-JUN-1999; 9905-0139463.
 PR 18-JUN-1999; 9905-0139750.
 PR 18-JUN-1999; 9905-0139763.
 PR 21-JUN-1999; 9905-0139817.
 PR 22-JUN-1999; 9905-0139899.
 PR 23-JUN-1999; 9905-0140353.
 PR 23-JUN-1999; 9905-0140354.
 PR 24-JUN-1999; 9905-0140695.
 PR 28-JUN-1999; 9905-0140823.
 PR 29-JUN-1999; 9905-0140991.
 PR 30-JUN-1999; 9905-0141287.
 PR 01-JUL-1999; 9905-0141842.
 PR 01-JUL-1999; 9905-0142154.
 PR 02-JUL-1999; 9905-0142055.
 PR 06-JUL-1999; 9905-0142390.
 PR 08-JUL-1999; 9905-0142803.
 PR 09-JUL-1999; 9905-0142977.
 PR 12-JUL-1999; 9905-0143542.
 PR 13-JUL-1999; 9905-0143542.
 PR 14-JUL-1999; 9905-0143624.
 PR 15-JUL-1999; 9905-0144005.
 PR 16-JUL-1999; 9905-0144085.
 PR 16-JUL-1999; 9905-0144086.
 PR 19-JUL-1999; 9905-0144325.
 PR 19-JUL-1999; 9905-0144331.
 PR 19-JUL-1999; 9905-0144332.
 PR 19-JUL-1999; 9905-0144333.
 PR 19-JUL-1999; 9905-0144334.
 PR 19-JUL-1999; 9905-0144335.
 PR 20-JUL-1999; 9905-0144352.
 PR 20-JUL-1999; 9905-0144632.
 PR 20-JUL-1999; 9905-0144884.
 PR 21-JUL-1999; 9905-0144814.
 PR 21-JUL-1999; 9905-0145086.

PR 21-JUL-1999; 9905-0145088.
 PR 22-JUL-1999; 9905-0145085.
 PR 22-JUL-1999; 9905-0145087.
 PR 22-JUL-1999; 9905-0145089.
 PR 22-JUL-1999; 9905-0145192.
 PR 23-JUL-1999; 9905-0145145.
 PR 23-JUL-1999; 9905-0145218.
 PR 23-JUL-1999; 9905-0145224.
 PR 26-JUL-1999; 9905-0145276.
 PR 27-JUL-1999; 9905-0145913.
 PR 27-JUL-1999; 9905-0145918.
 PR 27-JUL-1999; 9905-0145919.
 PR 28-JUL-1999; 9905-0145951.
 PR 02-AUG-1999; 9905-0146386.
 PR 02-AUG-1999; 9905-0146388.
 PR 02-AUG-1999; 9905-0146389.
 PR 03-AUG-1999; 9905-0147038.
 PR 04-AUG-1999; 9905-0147204.
 PR 04-AUG-1999; 9905-0147302.
 PR 05-AUG-1999; 9905-0147302.
 PR 05-AUG-1999; 9905-0147192.
 PR 05-AUG-1999; 9905-0147260.
 PR 06-AUG-1999; 9905-0147303.
 PR 06-AUG-1999; 9905-0147416.
 PR 09-AUG-1999; 9905-0147493.
 PR 09-AUG-1999; 9905-0147935.
 PR 10-AUG-1999; 9905-0148171.
 PR 11-AUG-1999; 9905-0148319.
 PR 12-AUG-1999; 9905-0148341.
 PR 13-AUG-1999; 9905-0148565.
 PR 13-AUG-1999; 9905-0148684.
 PR 16-AUG-1999; 9905-0149368.
 PR 17-AUG-1999; 9905-0149175.
 PR 18-AUG-1999; 9905-0149426.
 PR 20-AUG-1999; 9905-0149722.
 PR 20-AUG-1999; 9905-0149923.
 PR 20-AUG-1999; 9905-0149929.
 PR 23-AUG-1999; 9905-0149902.
 PR 25-AUG-1999; 9905-0150566.
 PR 26-AUG-1999; 9905-0150884.
 PR 27-AUG-1999; 9905-0151065.
 PR 27-AUG-1999; 9905-0151066.
 PR 27-AUG-1999; 9905-0151080.
 PR 27-AUG-1999; 9905-0151303.
 PR 30-AUG-1999; 9905-0151303.
 PR 31-AUG-1999; 9905-0151438.
 PR 01-SEP-1999; 9905-0151930.
 PR 07-SEP-1999; 9905-0152363.
 PR 10-SEP-1999; 9905-0153070.
 PR 13-SEP-1999; 9905-0153758.
 PR 15-SEP-1999; 9905-0154018.
 PR 16-SEP-1999; 9905-0154039.
 PR 20-SEP-1999; 9905-0154779.
 PR 22-SEP-1999; 9905-0155139.
 PR 23-SEP-1999; 9905-0155486.
 PR 24-SEP-1999; 9905-0155659.
 PR 28-SEP-1999; 9905-0156458.
 PR 29-SEP-1999; 9905-0156596.
 PR 04-OCT-1999; 9905-0157117.
 PR 05-OCT-1999; 9905-0157753.
 PR 06-OCT-1999; 9905-0157865.
 PR 07-OCT-1999; 9905-0158029.
 PR 08-OCT-1999; 9905-0158232.
 PR 12-OCT-1999; 9905-0158369.
 PR 13-OCT-1999; 9905-0159293.
 PR 13-OCT-1999; 9905-0159294.
 PR 13-OCT-1999; 9905-0159295.
 PR 14-OCT-1999; 9905-0159329.
 PR 14-OCT-1999; 9905-0159330.
 PR 14-OCT-1999; 9905-0159331.
 PR 14-OCT-1999; 9905-0159637.
 PR 14-OCT-1999; 9905-0159638.
 PR 18-OCT-1999; 9905-0159584.
 PR 21-OCT-1999; 9905-0160741.

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 1792;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
DB 1063 CCTGATGCTAAGTAGAAC 1044

RESULT 11

C34601/c
ID C34601 standard; DNA: 1797 BP.

XX C34601;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 7222.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;:

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134376.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 1797;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ccagatgataaagaac 20
Db 1065 CCTGATGTCATAGTAGAAC 1046

RESULT 12
C52084/c
ID C52084 standard; DNA; 1892 BP.
XX
AC C52084;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 70346.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140699.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0148368.
PR 17-AUG-1999; 99US-0148175.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.

PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 1892;
 Best Local Similarity 85.0%; Pred. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
 ||||| ||| ||||| |||||
 Db 1051 CCAGAGCGTTTAAAGTTGAC 1032

RESULT 13

T61274
 ID T61274 standard; DNA; 2329 BP.

AC T61274;

DT 23-APR-1997 (first entry)

DE Penicillium decumbens epoxidation enzyme genomic DNA.

KW Epoxidase; epoxidation; antibiotic; phosphomycin; precursor;
 (2)-1-propenyl phosphonic acid; ds.

OS Penicillium decumbens.

PH Key Location/Qualifiers
 FT exon 1..1294
 FT /*tag= a
 FT /number= 1

FT Intron /codon.start= 1007
 FT 1295..1421

FT /*tag= b
 FT /number= 1

FT exon 1422..1854
 FT /*tag= c
 FT /number= 2

FT Intron 1855..1916
 FT /*tag= d
 FT /number= 2

FT exon 1917..2329
 FT /*tag= e

PN JP08322560-A.

PD 10-DEC-1996.

PF 02-JUN-1995; 95JP-0136384.

PR 02-JUN-1995; 95JP-0136384.

XX (MEIJ) MEIJI SEIKA KAISHA LTD.

XX WPI: 1997-081076/08.

XX P-PSDB; W11700.

XX Epoxidation enzyme gene from Penicillium decumbens - useful for
 PT increased production of phosphomycin

XX Claim 5; Page 7-9; 10pp; Japanese.

XX The present genomic sequence from Penicillium decumbens encodes an
 CC epoxidation enzyme. The enzyme converts the phosphomycin precursor
 CC (2)-1-propenyl phosphonic acid into the antibiotic phosphomycin. The
 CC gene can be inserted into host cells to increase epoxidation
 CC activity, e.g. by up to 4.3-fold.

SO Sequence 2329 BP; 597 A; 587 C; 623 G; 522 T; 0 other;

Query Match 76.0%; Score 15.2; DB 18; Length 2329;
 Best Local Similarity 85.0%; Pred. No. 1.4e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
 | ||||| ||| | |||||
 Db 1900 ctatagtggtacatgtagaac 1919

RESULT 14

O94253
 ID O94253 standard; DNA; 4780 BP.

AC O94253;

DT 10-MAY-1996 (first entry)

DE Neuronal nitrogen monoxide synthase DNA.

KW nitrogen monoxide synthase; neuronal; treatment; prevention;
 KW vascular disease; restenosis; ds.

OS Homo sapiens.

PH Key Location/Qualifiers
 FT CDS 431..4732
 FT /*tag= a
 FT /product= bNOS

PN DE4411402-A1.

XX 05-OCT-1995.

PF 31-MAR-1994; 94DE-4411402.

PR 31-MAR-1994; 94DE-4411402.

XX (SCHR/) SCHRADER J.

XX Goedecke A, Schrader J;

XX WPI: 1995-345550/45.

XX P-PSDB; R77362.

XX Eukaryotic expression vector for nitrogen-mon:oxide synthase gene
 PT useful in the treatment and prevention of diseases of blood vessels
 PT by gene therapy

XX Claim 5; Fig 2; 28pp; German.

XX The DNA encodes the neuronal (brain-derived) nitrogen monoxide synthase
 CC (bNOS) isolated from humans. bNOS is a homodimer with a mol. wt. of 160
 CC kDa per subunit. Whereas endothelial NOS (eNOS) (R77363) binds calmodulin
 CC and is dependent on free calcium concentration, this enzyme was only
 CC active when the intracellular calcium concentration was increased as a
 CC result of receptor-mediated calcium influx. Vectors contg. The DNA are
 CC used in the treatment or prevention of vascular diseases, high blood
 CC pressure, arteriosclerosis, stenosis or restenosis of blood vessels, esp.
 CC coronary vessels after percutane transluminal coronary angioplasty. See
 CC O94255 and O94255 for inducible and endothelial NOS.

XX Sequence 4780 BP; 1159 A; 1413 C; 1292 G; 916 T; 0 other;

Query Match 76.0%; Score 15.2; DB 16; Length 4780;
 Best Local Similarity 85.0%; Pred. No. 1.5e+02;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
 ||||| ||| ||||| |||||
 Db 3616 ccagatggtataagtagaac 3635


```

RESULT 15
V87826
ID V87826 standard; cDNA; 653 BP.
XX
AC V87826;
XX
DT 12-FEB-1999 (first entry)
XX
DE EST clone EW13.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN W03845437-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US06956.
XX
PR 10-APR-1997; 97US-0837312.
XX
PA (GEMV ) GENETICS INST INC.
XX
PI Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D;
PI Racie LA, Spaulding V, Treacy M;
XX
DR WPI: 1999-070078/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from
PT e.g. human blood, kidney, foetal lung, placenta, testes, brain,
PT ovary, pituitary, retina and colon cDNA libraries
XX
PS Claim 1; Page 192-193; 641pp; English.
XX
CC The present sequence represents an expressed sequence tag (EST), and is
CC a polynucleotide of the invention. The polynucleotides of the invention
CC are all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene
CC therapy.
XX
SQ Sequence 653 BP; 210 A; 108 C; 147 G; 188 T; 0 other;

```

```

Query Match 75.0%; Score 15; DB 20; Length 653;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 tggataagtagaac 20
| | | | | | | | | | |
DB 557 tggataagtagaac 571

```

Search completed: June 8, 2001, 06:21:25
 Job time: 6149 sec

.....

This Page Blank (uspto)

Source

```

source
1. 313
   /organism="Mus musculus"
   /strain="C57BL/6J"
   /db_xref="taxon:10090"
   /clone_1b="IMAGE:636532"
   /clone_1lb="Soares mouse lymph node NbM1N"
   /sex="male"
   /tissue_type="lymph node"
   /dev_stage="4 weeks"
   /lab_host="DH10B"
   /note="Organ: Lymph node; Vector: pT7n3D-Pac (Pharmacia)
   /site_1: Not I; Site_2: Eco RI;
   1st strand cDNA was primed with a Not I - oligo(dT) primer
   15'
   TGTACCACTCATGAGTGGAGCGCCGCGACACTCTTTTCTTTTTTTTTTTTTTTT
   3'}. double-stranded cDNA was ligated to Eco RI adaptors
   (Pharmacia), digested with Not I and cloned into the Not I
   and Eco RI sites of the modified pT73 vector. RNA
   provided by Dr. Bertrand Jordan. Library constructed and
   normalized by Bento Soares and M. Fatima Bonaldo."
83 a      83 c      79 g      68 t

```

		79.0%;	Score 15.8;	DB 23;	Length 313;
Query Match			Pred. No. 2.9e+02;		
Best Local Similarity	89.5%;		Mismatches	Indels 0;	Gaps 0;
Matches 17;	Conservative	0;			
QY	2 aatgacgggaactctcg	20			
db	233 AGTACCGGAAACTCTTCG	251			

RESULT	11			
W38506/c				
LOCUS	334 bp	EST	15-MAY-1996	
DEFINITION	W38506			
	ZD19605.r1 Soares.fetal_lung NbHL19W Homo sapiens	CDNA clone		
	IMAGE:302528 5', mRNA sequence.			

ACCESSION	W38506	
VERSION	W38506.1	GI:1320111
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
		Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANISM	REFERENCE
Human sapiens	1 (bases 1 to 334)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
Hillier, L., Clark, N., Dubugne, T., Ellison, K., Hawkins, M., Holman	
W., Jones, A., Krawinkel, M., Marshall, M., Miller, C., Murrells, J.,	
Paterson, G., Pearson, M., Rieu, M., Searles, B., Smith, J., Steward	
et al., 1996	

TITLE M.M. Hultman, M. Aucca, E. Lehtinen, R. Rifkin, L. Rohling, T. Soares, M. Tan, F. Trevarski, E. Waters, R. Williamson, A. Wohlmann, P. and Wilson, R.
The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine

Contact: Washington University School of Medicine
 660 South Euclid Avenue, Box 8501, St. Louis, MO 63108
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel.: 314 286 1800
 Fax: 314 286 1810
 Email: es@watson.wustl.edu
 This clone is available royalty-free through LIND, contact the
 IMAGE Consortium (info@image.lind.gov) for further information.
 Seq primer: mob.REGA+EI
 High quality sequence stop: 268.

FEATURES
source

```

1..334
/organism="Homo sapiens"
/db.xref="GDB:1247460"
/db.xref="taxon:9606"
/clone="IMAGE:302528"
/clone_1lb="Soares_fetal_lung_NBHL19w"
/dev_host="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site:1: Not I; Site,2: Eco RI; 1st

```

	60 a	79 c	114 g	79 t	2 others
BASE COUNT					
ORIGIN					

Query Match	79.08;	Score 15.8;	DB 158;	Length 334;
Best Local Similarity	85.08;	Freqd No. 3e+02;	3;	Indels 0;
Matches 17;	Conservative	0;	Mismatches	Gaps
QY 1 caatgaacgggaacattccg 20				
DB 311 cactatcagcaggaacattccg 292				

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL
I (Passes 1 to 30)	Marrin, M., Hillier, L., Allen, M., Bowles, N., Dierich, N., Dubouquet, T.	The Washu-HNMI Mouse EST Project	unpublished (1996)
	Gisellel, S., Kucaba, T., Lacey, M., Le M., Matlin, J., Morris, M., Schellensberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Welling, B., Wyllie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.	The Washu-HNMI Mouse EST Project	unpublished (1996)
		Contact: Marria M/Mouse Est Project	
		Washu-HNMI Mouse EST Project	

Washington University School of Medicine
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: moussese@atson.wustl.edu
This clone is available royalty-free through LLNL / contact the
IMAB Consortium (info@imab.llnl.gov) for further information.
MGI:388524
seq primer: -28M13 rev2 from Amersham.

FEATURES
source

```

Location/Qualifiers
1..350
/organism="Mus musculus"
/strain="C57Bl/6J"
/db_xref="taxon:10090"
/clone="IMAGE:636533"
/clone_id="Scars mouse lymph node N8MLN"
/sex="male"
/tissue_type="lymph node"
/dev_stage="4 weeks"
/lab_host="DH10B"
note="Organ: lymph node; Vector: pT73d-Pac (Pharmacia)
with a modified polylinker. Site.1: Not I; Site.2: Eco RI;
1st strand cDNA was primed with a Not I - oligo(dT) primer
15'
TGTACCACTCTGAAGCGAGGCGGCGGATCACTTTTTTTTTTTTTTTTTT
3'; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. RNA

```

This Page Blank (uspto)

provided by Dr. Bertrand Jordan. Library constructed and
normalized by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 118 a 82 c 81 g 69 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 3; Length 350;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Yy 2 aatgacgggaacgttcg 20
Db 234 AGTACCGGAACGTTCCG 252

RESULT 13
LOCUS C69306 378 bp mRNA EST 23-SEP-1997
DEFINITION C69306 Yuj1 Kohara unpublished cDNA Caenorhabditis elegans CDNA
clone YK361c9 5', mRNA sequence.
ACCESSION C69306
VERSION C69306.1 GI:2430662
KEYWORDS EST
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
; Rhabditidae; Pelodidae; Caenorhabditis.
1 (bases 1 to 378)
Kohara, Y., Motobashi, T., Tabara, H., Watanabe, H., Sugimoto, A., Sano
M., Miyata, A. and Nishigaki, A.
Expression map of the C. elegans genome
Unpublished (1996)
Contact: Yuj1 Kohara
Gene Library Lab
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
source
FEATURES
location/Qualifiers
1. 378
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"
/clone="YK361c9"
/clone_1b="Yuj1 Kohara unpublished CDNA"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
/dev_stage="varied"

BASE COUNT 103 a 94 c 75 g 106 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 175; Length 378;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Yy 2 aatgacgggaacgttcg 20
Db 75 AATCAGGGAACGTTCCG 93

RESULT 14
LOCUS A1026675 420 bp mRNA EST 07-JUL-1998
DEFINITION A1026675 x1 Soares, testis_NHR Homo sapiens CDNA clone IMAGE:1645646
3', mRNA sequence.
ACCESSION A1026675
VERSION A1026675.1 GI:3246163
KEYWORDS EST
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 420)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo
Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/bdrrp/image/image.html
Insert length: 529 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 399.
location/Qualifiers
1. 420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1645646"
/clone_1b="Soares-testis_NHR"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand CDNA
was prepared from mRNA obtained from clonech laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5']
TCTTACCATCTGACAGGAGGCGGCCCAATTTTITTTTITTTT 3']
Double-stranded CDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 113 a 103 c 105 g 99 t
ORIGIN

Query Match 79.0%; Score 15.8; DB 15; Length 420;
Best Local Similarity 89.5%; Pred. No. 3.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Yy 2 aatgacgggaacgttcg 20
Db 266 AATGATGGAACGTTCCG 284

RESULT 15
LOCUS AL381271 502 bp mRNA EST 03-AUG-2000
DEFINITION M16C019A02P1 M16C Medicago truncatula CDNA clone M16C019A0 T3, mRNA
sequence.
ACCESSION AL381271
VERSION AL381271.1 GI:9681022
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 502)
Journet, E.P., Crespeau, H., Van Tuinen, D., Gouzy, J., Jallion, O.,
Miebel, A., Carreau, V., Chataigner, O., Kahn, D., Glanina, Z., Pearson
V., and Gamas, P.
Medicago truncatula ESTs from endomycorrhizal roots
Unpublished (2000)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr

This Page Blank (uspto)

misc
misc-1
misc-fea
misc-fea

Query Match 84.0%; Score 16.8; DB 87; Length 220922;
Best Local Similarity 90.0%; Pred. No. 1.9e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagatgataagtagaac 20
DB 127821 CCAGATGATTAATAATAGAAC 127802

RESULT 11
RNCDDAK 989 bp mRNA ROD 04-JAN-1995
LOCUS R.norvegicus mRNA for Cdk-activating kinase.
DEFINITION X83579
ACCESSION X83579.1 GI:619508
VERSION X83579.1 GI:619508
KEYWORDS Cdk-activating kinase.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 989)
AUTHORS Wu, L. and Hall, F.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 989)
AUTHORS Hall, F.
TITLE Direct Submission
COMMENT Submitted (19-DEC-1994) F. Hall, Children's Hospital Los Angeles,
Division of Orthopaedic Surgery, 4650 Sunset Boulevard, Los
Angeles, CA 90027, USA
LOCATION/Qualifiers
FEATURES
source 1. .989
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="testis"
/cell_type="mixed"
/cell_line="rat testis CDNA"
1. .941
/gene="CAK1"
/codon_start=1
/product="Cdk-activating kinase Cdk7"
/protein_id="CA58562.1"
/db_xref="GI:619509"
/db_xref="SWISS-PROT:P51952"
/translation="ANNREKIDPLGEGOPATYKARKDNTNOIVAIIKKIKLGRSEAK
DGINKTALREIKLQELSHPTIGLDAGFKHNSISLVDPMEYDLEVIITKDNLSVLT
PSHAKVAMIMTLOGLIEYLHOMWILHRDLKNNLLDENGVKLADFGSLAKSGSNNA
YTHOVYTRMYRARELLFGARMYGVGVMMVAVGCIILAEILLRVPFLPGSDLDLQTRIF
ETLGHPTDEQMDKSLPDIYTKSPGIPLOHIFIAAGDDLELLIQLGFLFNPCTRI
TASQLARKYFSNRPGTPGQQLPRPCVPEALKEQ"

BASE COUNT 291 a 217 c 235 g 246 t
ORIGIN
Query Match 82.0%; Score 16.4; DB 94; Length 989;
Best Local Similarity 94.4%; Pred. No. 3.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 agatgataagtagaac 20
DB 134 AGATGATTAATAATAGAAC 151
RESULT 12
HSCAKIMR 1269 bp mRNA PRI 08-AUG-1996
LOCUS HSCAKIMR
DEFINITION H.sapiens CAK1 mRNA for Cdk-activating kinase.
ACCESSION X77303
VERSION X77303.1 GI:1486356

KEYWORDS Cdk-activating kinase; protein kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1269)
AUTHORS Wu, L., Yee, A., Liu, L., Carbonaro-Hall, D., Venkatesan, N., Tolo, V. T.
and Hall, F. L.
TITLE Molecular cloning of the human CAK1 gene encoding a
cyclin-dependent kinase-activating kinase
JOURNAL Oncogene 9 (7), 2089-2096 (1994)
MEDLINE 94268852
REFERENCE 2 (bases 1 to 1269)
AUTHORS Hall, F. L.
TITLE Direct Submission
COMMENT Submitted (14-JAN-1994) F. L. Hall, Childrens Hospital Los Angeles,
Division of Orthopaedic Surgery, 4650 Sunset Boulevard, Los
Angeles, California 90027, USA
LOCATION/Qualifiers

REMARK 3 (bases 1 to 1269)
Wu, L.
AUTHORS Direct Submission
TITLE Submitted (08-AUG-1996) L. Wu, Childrens Hospital Los Angeles,
Division of Orthopaedic Surgery, 4650 Sunset Boulevard, Los
Angeles, California 90027, USA
JOURNAL On Aug 9, 1996 this sequence version replaced gi:509772.
COMMENT
FEATURES
source 1. .1269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="connective"
/cell_type="fibroblast"
/cell_line="W138 human fibroblast (ATCC)"
/clone_1b="lambda CDNA"
/clone="pBluescript Phagemid"
44. .1084
/gene="CAK1"
44. .1084
/gene="CAK1"
/codon_start=1
/product="Cdk-activating kinase"
/protein_id="CA54508.1"
/db_xref="GI:1486357"
/db_xref="SWISS-PROT:P50613"
/translation="VALDYKSRKRYEKIDELGEGOPATYKARKDNTNOIVAIIKKI
LGHSERKDGINTALREIKLQELSHPTIGLDAGFKHNSISLVDPMEYDLEVIIT
KNSLVTPSHIRAYVIMTLOGLIEYLHOMWILHRDLKNNLLDENGVKLADFGSLAK
SGSPNRYTHOVYTRMYRARELLFGARMYGVGVMMVAVGCIILAEILLRVPFLPGSD
LDLQTRIFETLGHPTDEQMDKSLPDIYTKSPGIPLOHIFIAAGDDLELLIQLGFL
FNPCTRI TASQLARKYFSNRPGTPGQQLPRPCVPEALKEQSNALAIKKRTEA
LEOGLPKRLIF"

BASE COUNT 425 a 227 c 277 g 340 t
ORIGIN
Query Match 82.0%; Score 16.4; DB 92; Length 1269;
Best Local Similarity 94.4%; Pred. No. 3.7e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 agatgataagtagaac 20
DB 199 AGATGATTAATAATAGAAC 216
RESULT 13
HSSTPMO15 1274 bp mRNA PRI 27-MAY-1997
LOCUS HSSTPMO15
DEFINITION H.sapiens mRNA for serine/threonine protein kinase MO15.
ACCESSION Y13120
VERSION Y13120.1 GI:2125815
KEYWORDS serine/threonine protein kinase.
SOURCE human.
ORGANISM Homo sapiens

This Page Blank (uspto)

This Page Blank (uspto)

Fri Jun 8 09:43:27 2001

us-09-507-242-2.rag

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 8, 2001, 09:37:50 ; Search time 20.01 Seconds
(without alignments)
505.642 Million cell updates/sec

Title: US-09-507-242-2

Perfect score: 932
Sequence: 1 MASTSTGLPPNMTIRVSRSH.....GEVSNITETSGVHLORTG 177

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A-Geneseg.0401:*
2: /SIDS6/gcgdata/geneseg/genesegp/AA1981.DAT:*
3: /SIDS6/gcgdata/geneseg/genesegp/AA1982.DAT:*
4: /SIDS6/gcgdata/geneseg/genesegp/AA1983.DAT:*
5: /SIDS6/gcgdata/geneseg/genesegp/AA1984.DAT:*
6: /SIDS6/gcgdata/geneseg/genesegp/AA1985.DAT:*
7: /SIDS6/gcgdata/geneseg/genesegp/AA1986.DAT:*
8: /SIDS6/gcgdata/geneseg/genesegp/AA1987.DAT:*
9: /SIDS6/gcgdata/geneseg/genesegp/AA1988.DAT:*
10: /SIDS6/gcgdata/geneseg/genesegp/AA1989.DAT:*
11: /SIDS6/gcgdata/geneseg/genesegp/AA1990.DAT:*
12: /SIDS6/gcgdata/geneseg/genesegp/AA1991.DAT:*
13: /SIDS6/gcgdata/geneseg/genesegp/AA1992.DAT:*
14: /SIDS6/gcgdata/geneseg/genesegp/AA1993.DAT:*
15: /SIDS6/gcgdata/geneseg/genesegp/AA1994.DAT:*
16: /SIDS6/gcgdata/geneseg/genesegp/AA1995.DAT:*
17: /SIDS6/gcgdata/geneseg/genesegp/AA1996.DAT:*
18: /SIDS6/gcgdata/geneseg/genesegp/AA1997.DAT:*
19: /SIDS6/gcgdata/geneseg/genesegp/AA1998.DAT:*
20: /SIDS6/gcgdata/geneseg/genesegp/AA1999.DAT:*
21: /SIDS6/gcgdata/geneseg/genesegp/AA2000.DAT:*
22: /SIDS6/gcgdata/geneseg/genesegp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	932	100.0	177	21	Y95876
2	369.5	39.6	163	18	W18312
3	313.5	33.6	163	20	Y48377
4	303.5	32.6	119	21	G06627
5	303.5	32.6	119	21	G14662
6	303.5	32.6	119	21	G43433
7	157	16.8	92	21	Y90945
8	142.5	15.3	320	18	W89744
9	137	14.7	221	21	G08943
10	137	14.7	299	21	G08942
11	137	14.7	303	21	G08941

12	127.5	13.7	142	21	G11000
13	127.5	13.7	169	21	G10999
14	121.5	13.0	204	20	W89814
15	121.5	13.0	299	18	W55450
16	121.5	13.0	299	19	W98322
17	121.5	13.0	299	19	W73034
18	121.5	13.0	299	20	W89829
19	121.5	13.0	299	20	W89849
20	121.5	13.0	39	21	B21976
21	117	12.6	37	17	R97692
22	116.5	12.5	317	20	Y19968
23	116.5	12.5	336	20	Y19968
24	102	10.9	131	19	W9203
25	101.5	10.8	138	20	G08064
26	101	10.2	115	21	G34784
27	95.5	10.2	143	21	G34784
28	95.5	10.2	143	21	G34784
29	93	10.0	39	21	B21975
30	93	10.0	31	21	B21943
31	92	9.9	31	21	B21943
32	86.5	9.3	428	20	Y09000
33	85	9.1	1584	21	Y84430
34	85	9.1	1693	21	B48457
35	85	9.1	1693	21	B48457
36	85	9.1	1713	21	B48457
37	85	9.1	1724	21	B48457
38	84.5	9.1	877	19	W98214
39	84	9.0	331	21	G05720
40	84	9.0	638	21	G38502
41	84	9.0	720	21	G38501
42	84	9.0	777	21	G38500
43	83.5	8.9	286	18	W55712
44	82.5	8.9	62	8	P70035
45	82	8.8	326	22	B60765

ALIGNMENTS

RESULT 1	
Y95876	standard; Protein; 177 AA.
AC	Y95876;
XX	20-NOV-2000 (first entry)
XX	
XX	Candida albicans Caes1 proteins.
XX	Caes1; Ess1; infection; diagnosis; therapy; antifungal; fungicide;
XX	antiproliferative; cytostatic; antitumour.
KW	
KW	Candida albicans.
XX	
OS	
XX	WO200050561-A2.
PN	
XX	31-AUG-2000.
PD	
XX	18-FEB-2000; 2000MO-US04203.
PF	
XX	23-FEB-1999; 9905-0121246.
PR	
XX	(HEAL-) HEALTH RES INC.
PA	
XX	Hanes SD, Devasahayam G, Chaturvedi V;
PI	
XX	WPI. 2000-565453/52.
DR	N-PSDB; A50292.
XX	Novel Candida albicans gene, Caes1 useful for identifying compounds
PT	that specifically bind to and/or inhibit Caes1 and thus for treating
PT	Candida albicans infections and other life-threatening fungal
PT	infections -

Arabidopsis thalia
Arabidopsis thalia
Protein encoded by
H. pylori ORF 02ae
H. pylori GHPO 136
Helicobacter pylori
Protein encoded by
Antigen from clust
ESS1/9C peptide co
Mouse Ess1 WW doma
B. burgdorferi ant
Human parvulin-11k
S. typhimurium sur
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
Ess1/S.c. peptide
Prl1/human peptide
E. coli sura prote
Human laminin 5 po
Human laminin 5 po
Human laminin 108
H. pylori GHPO 108
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
H. pylori ORF 06gp
Secretory signal s
Gene 16 related pe

XX Claim 4; Fig 1a; S1pp: English.
 PS This is the predicted translation product, Caes1, of the Caes1
 XX gene from *Candida albicans*. Caes1 is the C. albicans homologue of
 CC Saccharomyces cerevisiae Ess1, showing 42% amino acid identity.
 CC Caes1 protein is a target for screening for antifungal and
 CC inhibitor compounds useful for treating or preventing C. albicans
 CC infections, and for raising antibodies useful for diagnostic
 CC purposes or for blocking Caes1 activity. A Caes1 inhibitor can
 CC be a compound which selectively inhibits growth of S. cerevisiae
 CC not containing an endogenous Ess1 gene but rather Caes1 and
 CC uninduced P1n1 (a human homologue of Ess1) and/or preferably does
 CC not inhibit induced P1n1, e.g. does not inhibit S. cerevisiae not
 CC containing an endogenous Ess1 but rather induced P1n1. Compositions
 CC which inhibit P1n1 are useful antiproliferatives e.g. antineoplastics,
 CC anticancer agents or anticancer agents. An anti-Caes1 antibody
 CC or an antibody against an epitopic region of Caes1 can also be an
 CC inhibitor of Caes1 by virtue of the antibody being able to bind to
 CC Caes1.
 XX
 XX Sequence 177 AA:
 SQ
 Query Match 100.0%; Score 932; DB 21; Length 177;
 Best Local Similarity 100.0%; Pred. No. 2.4e-90;
 Matches 177; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MASSTGLPFWTIRVSRSHNKEFLNOSTNESSWDPPYGTDEKVLNAYIAKFRNNGKPLVNE
 DB 1 masstglpfwtlrvsrshnkeflnqstnesswdppygtdkavlakikmngykp 60
 QY 61 LVNEDGVVRVSHLIRKNNQSRKPKSWSPDGISRTDESIOLKKHLERILSGEVKLSDEL 120
 DB 61 lvnedgvvrshllirknnqsrkpkswspdgistrdesiolkkhlerrilsgevksel 120
 QY 121 ANTESDCSSHDGRGDLGFESKQOMOPPEEAFFNLHGEVSNIIETNSGVHILORTG 177
 DB 121 antesdcsshdrgdglgfeskgomoppfeeaafnlhgevsniietnsgvhlgrtg 177
 RESULT 2
 ID W18312 standard; Protein; 163 AA.
 XX
 AC W18312;
 XX
 DT 09-DEC-1997 (first entry)
 XX
 DE NIMA-interacting protein P1n1.
 XX
 KW P1n1; protein interacting with NIMA; cell proliferation; mitosis;
 KM peptidyl-propyl cis/trans isomerase; adenocarcinoma; cancer;
 KM leukaemia; psoriasis; pemphigus vulgaris; rheumatoid arthritis;
 KW acute respiratory distress syndrome; septic shock; inflammation;
 KW therapy.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Domain 4..43
 FT /label= MW
 FT /note= "conserved tryptophan domain thought to
 FT be mediate protein-protein interactions"
 FT Peptide 54..69
 FT /note= "putative nuclear localisation signal"
 FT Domain 59..163
 FT /label= PPI
 FT /note= "peptidyl-propyl cis/trans isomerase domain"
 XX MO9717986-A1.
 XX 22-MAY-1997.
 PD

XX
 PF 28-OCT-1996; 96MO-US17334.
 XX
 PR 13-NOV-1995; 95US-0555912.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Hunter T, Lu KP;
 XX WPI: 1997-289057/26.
 DR N-PSDB; T68888.
 XX
 XX Protein, P1n1, interacting with NIMA - used for treating cell
 XX proliferative disorders
 PS Claim 4; Page 49-50; 73pp: English.
 CC This human polypeptide sequence comprises P1n1, an 18 kDa protein
 CC that has peptidyl-propyl cis/trans isomerase activity, associates
 CC with NIMA protein kinase, inhibits the mitosis-promoting function
 CC of NIMA when overexpressed, and induces mitotic arrest and nuclear
 CC fragmentation when depleted. Its sequence was deduced from a DNA
 CC sequence (T68888) identified in an HeLa library using a yeast two-
 CC hybrid system. A recombinant expression vector comprising the
 CC DNA sequence and host cells containing the vector are claimed.
 CC Methods are also claimed for identifying proteins that inhibit the
 CC mitosis promoting function of NIMA protein kinase and for
 CC controlling the growth of a cell by reducing P1n1 activity or P1n1
 CC expression using an inhibitor, anti-P1n1 antibody, antisense
 CC nucleotide sequence or ribozyme, or by increasing P1n1 activity
 CC in the presence of an activator or increasing P1n1 expression using
 CC an enhancer. This allows treatment of cell proliferation disorders
 CC such as adenocarcinoma, cancer, psoriasis, pemphigus vulgaris, acute
 CC respiratory distress syndrome, rheumatoid arthritis, septic shock
 CC and inflammation.
 XX
 SQ Sequence 163 AA:
 QY
 Query Match 39.6%; Score 369.5; DB 18; Length 163;
 Best Local Similarity 43.2%; Pred. No. 3.8e-31;
 Matches 73; Conservative 34; Mismatches 49; Indels 13; Gaps 2;
 QY 8 LPPMWTIRVSRSHNKEFLNOSTNESSWDPPYGTDEKVLNAYIAKFRNNGKPLVNE
 DB 7 lppgwkrmsrsgryvyyfnhtltaasqwerpsgn-----ssgkngqgepar 54
 QY 68 VRVSHLIRKNNQSRKPKSWSPDGISRTDESIOLKKHLERILSGEVKLSDELSDC 127
 DB 55 vrcshllrvksgrsrrpswrq-ekltrckealelmgylqkiksgedfslasqfsc 113
 QY 128 SSHRGGDLGFESKQOMOPPEEAFFNLHGEVSNIIETNSGVHILORT 176
 DB 114 ssakarqdlgrsrgmqkprcdastalrtgmsgpvftdsghililrt 162
 RESULT 3
 ID Y48377 standard; Protein; 163 AA.
 XX
 AC Y48377;
 XX
 DT 08-DEC-1999 (first entry)
 XX
 DE Human prostate cancer-associated protein 74.
 XX
 KW Expressed sequence tag; EST; prostate; tumor; treatment; gene therapy;
 KW cancer; tissue specificity; human.
 XX
 OS Homo sapiens.
 XX
 XX Y48377
 XX DEL9811194-A1.
 XX

```
PD      16-SEP-1999.
XX      10-MAR-1998;          98DE-1011194.
PF      XX
XX      10-MAR-1998;          98DE-1011194.
PR      XX
PA      (META-) METAGEN GES GENOMFORSCHUNG MBH.
PI      Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
XX      WP1: 1999-519629/44.
DR      N-P5DB; Z33510.
XX      PT
PT       New nucleic acid expressed at high level in normal prostatic tissue and
        encoded polypeptides, used to treat cancer and screen for therapeutic
        agents -
XX      PS
PS     Claim 25; 152; 194pp; German.
XX      CC
XX      This invention describes novel nucleic acid sequences (A) that are
CC      expressed at high level in normal prostatic tissue. Polypeptides (I)
CC      encoded by (A) are used: (a) for identifying agents for treatment of
CC      prostatic cancer and (b) for therapy of prostate cancer, optionally
CC      where expressed by gene therapy methods. (A) is also used to isolate
CC      full-length genes (for gene therapy) and for recombinant production of
CC      (I), which can be used to raise specific antibodies. (A) are identified
CC      by assembly of ESTs (expressed sequence tags) before these are analyzed
CC      for expression pattern (tissue specificity). This approach eliminates
CC      many of the false results, as regards tissue specificity associated with
CC      known methods that use single (usually short) ESTs.: Y48304-Y48456
CC      represent peptides encoded by the expressed sequence tags described in
CC      the method of the invention.
XX      SQ
SQ     Sequence    163 AA;

Query Match           33.6%; Score 313.5; DB 20; Length 163;
Best Local Similarity 42.5%; Pred. No. 3e-25; Indels 13; Gaps 2
Matches   62; Conservative % 29; Mismatches 42;

QY      8 LPPNNTIRVSRSHNKNEYFLNSTNSSMDPEYGTDFKEVLNAVIAKFRNGYKPLVNEDQG 67
         ||| :||| | :| | | | | | | | | | | | | | | | | | | | | | :| | | :
Db       7 lppgyekmtsrsgvvyfnnhtnagwerpsgn-----ssggknqgsepar 54

QY      68 VVRSSHLLIKNNOSRRPKSNKSPDGISTRDESIGLKKHLERILISGEVKLSLANTESDC 127
         |||||||:|||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       55 vrrshlllvhsgrsrrpswrg-ektlrkeaelinglylqlkvsgeedfesiasqfsdc 113

QY      128 SSHDCGDLGFSGSKGMOPFEAAAF 153.
         || | || | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       114 ssakargdlgaifergmqkfedaast 139

RESULT   4
G06627 standard: Proteio; 119 AA.
AC      G06627;
XX      DE
XX      17-OCT-2000 (first entry)
XX      Arbidlopsis thaliana protein fragment SEQ ID NO: 3468.
XX      Protein identification: signal transduction pathway; metabolic pathway;
XX      hydrialsation assay; genetic mapping; gene expression control; promoter;
XX      termination sequence.
OS      Arabidopsis thaliana.
PN      EP1033405-A2.
XX      PD
PD     06-SEP-2000;
```

PR	06-JUL-1999;	99US-0142803
PR	02-JUL-1999;	99US-0142955
PR	01-JUL-1999;	99US-0142154
PR	30-JUN-1999;	99US-0141842
PR	29-JUN-1999;	99US-0141287
PR	28-JUN-1999;	99US-0140991
PR	24-JUN-1999;	99US-0140823
PR	24-JUN-1999;	99US-0140695
PR	23-JUN-1999;	99US-0140354
PR	22-JUN-1999;	99US-0140353
PR	21-JUN-1999;	99US-0139899
PR	18-JUN-1999;	99US-0138766
PR	18-JUN-1999;	99US-0138461
PR	18-JUN-1999;	99US-0138463
PR	18-JUN-1999;	99US-0138465
PR	18-JUN-1999;	99US-0138466
PR	18-JUN-1999;	99US-0138467
PR	18-JUN-1999;	99US-0139458
PR	18-JUN-1999;	99US-0139459
PR	18-JUN-1999;	99US-0139460
PR	18-JUN-1999;	99US-0139461
PR	18-JUN-1999;	99US-0139462
PR	17-JUN-1999;	99US-0139442
PR	16-JUN-1999;	99US-0139453
PR	15-JUN-1999;	99US-0139443
PR	14-JUN-1999;	99US-0139119
PR	10-JUN-1999;	99US-0138847
PR	08-JUN-1999;	99US-0138540
PR	07-JUN-1999;	99US-0137724
PR	04-JUN-1999;	99US-0137502
PR	03-JUN-1999;	99US-0137528
PR	28-MAY-1999;	99US-0136782
PR	27-MAY-1999;	99US-0136592
PR	22-MAY-1999;	99US-0136321
PR	21-MAY-1999;	99US-0136029
PR	20-MAY-1999;	99US-0135524
PR	19-MAY-1999;	99US-0134941
PR	18-MAY-1999;	99US-0134768
PR	14-MAY-1999;	99US-0134370
PR	14-MAY-1999;	99US-0134321
PR	14-MAY-1999;	99US-0134219
PR	11-MAY-1999;	99US-0134218
PR	07-MAY-1999;	99US-0134256
PR	06-MAY-1999;	99US-0132487
PR	05-MAY-1999;	99US-0132485
PR	04-MAY-1999;	99US-0132484
PR	30-APR-1999;	99US-0132407
PR	28-APR-1999;	99US-0132408
PR	23-APR-1999;	99US-0131449
PR	21-APR-1999;	99US-0130991
PR	19-APR-1999;	99US-0130449
PR	16-APR-1999;	99US-0129845
PR	08-APR-1999;	99US-0128734
PR	06-APR-1999;	99US-0126762
PR	29-MAR-1999;	99US-0126785
PR	25-MAR-1999;	99US-0126264
PR	05-MAR-1999;	99US-0125548
PR	05-MAR-1999;	99US-0123180
PR	25-FEB-1999;	99US-0121825
PR	25-FEB-2000;	2000EP-0301439
PR		

PR	09-JUL-1999;	PR	99US-0142920.
PR	12-JUL-1999;	PR	99US-0142977.
PR	13-JUL-1999;	PR	99US-0143542.
PR	14-JUL-1999;	PR	99US-0143624.
PR	15-JUL-1999;	PR	99US-0144005.
PR	16-JUL-1999;	PR	99US-0144085.
PR	16-JUL-1999;	PR	99US-0144086.
PR	19-JUL-1999;	PR	99US-0144325.
PR	19-JUL-1999;	PR	99US-0144331.
PR	19-JUL-1999;	PR	99US-0144332.
PR	19-JUL-1999;	PR	99US-0144333.
PR	19-JUL-1999;	PR	99US-0144334.
PR	20-JUL-1999;	PR	99US-0144335.
PR	20-JUL-1999;	PR	99US-0144652.
PR	20-JUL-1999;	PR	99US-0144884.
PR	21-JUL-1999;	PR	99US-0144814.
PR	21-JUL-1999;	PR	99US-0145086.
PR	21-JUL-1999;	PR	99US-0145087.
PR	22-JUL-1999;	PR	99US-0145085.
PR	22-JUL-1999;	PR	99US-0145089.
PR	22-JUL-1999;	PR	99US-0145192.
PR	23-JUL-1999;	PR	99US-0145145.
PR	23-JUL-1999;	PR	99US-0145218.
PR	23-JUL-1999;	PR	99US-0145224.
PR	26-JUL-1999;	PR	99US-0145276.
PR	27-JUL-1999;	PR	99US-0145913.
PR	27-JUL-1999;	PR	99US-0145918.
PR	27-JUL-1999;	PR	99US-0145919.
PR	28-JUL-1999;	PR	99US-0145951.
PR	02-AUG-1999;	PR	99US-0146386.
PR	02-AUG-1999;	PR	99US-0146388.
PR	02-AUG-1999;	PR	99US-0146389.
PR	03-AUG-1999;	PR	99US-0147038.
PR	04-AUG-1999;	PR	99US-0147204.
PR	04-AUG-1999;	PR	99US-0147302.
PR	05-AUG-1999;	PR	99US-0147192.
PR	05-AUG-1999;	PR	99US-0147260.
PR	06-AUG-1999;	PR	99US-0147303.
PR	06-AUG-1999;	PR	99US-0147416.
PR	09-AUG-1999;	PR	99US-0147493.
PR	09-AUG-1999;	PR	99US-0147935.
PR	10-AUG-1999;	PR	99US-0148171.
PR	11-AUG-1999;	PR	99US-0148319.
PR	12-AUG-1999;	PR	99US-0148341.
PR	13-AUG-1999;	PR	99US-0148565.
PR	13-AUG-1999;	PR	99US-0148684.
PR	16-AUG-1999;	PR	99US-0149368.
PR	17-AUG-1999;	PR	99US-0149175.
PR	18-AUG-1999;	PR	99US-0149426.
PR	20-AUG-1999;	PR	99US-0149722.
PR	20-AUG-1999;	PR	99US-0149723.
PR	20-AUG-1999;	PR	99US-0149929.
PR	23-AUG-1999;	PR	99US-0149902.
PR	23-AUG-1999;	PR	99US-0149930.
PR	25-AUG-1999;	PR	99US-0150566.
PR	26-AUG-1999;	PR	99US-0150884.
PR	27-AUG-1999;	PR	99US-0151065.
PR	27-AUG-1999;	PR	99US-0151066.
PR	27-AUG-1999;	PR	99US-0151080.
PR	30-AUG-1999;	PR	99US-0151303.
PR	31-AUG-1999;	PR	99US-0151348.
PR	01-SEP-1999;	PR	99US-0151930.
PR	07-SEP-1999;	PR	99US-0152363.
PR	10-SEP-1999;	PR	99US-0153070.
PR	13-SEP-1999;	PR	99US-0153758.
PR	15-SEP-1999;	PR	99US-0154018.
PR	16-SEP-1999;	PR	99US-0154039.
PR	20-SEP-1999;	PR	99US-0154779.
PR	22-SEP-1999;	PR	99US-0155139.
PR	23-SEP-1999;	PR	99US-0155486.
PR	24-SEP-1999;	PR	99US-0155659.
PR	28-SEP-1999;	PR	99US-0156458.
PR	29-SEP-1999;	PR	99US-0156596.
PR	04-OCT-1999;	PR	99US-0157117.
PR	05-OCT-1999;	PR	99US-0157753.
PR	06-OCT-1999;	PR	99US-0157865.
PR	07-OCT-1999;	PR	99US-0158029.
PR	08-OCT-1999;	PR	99US-0158232.
PR	12-OCT-1999;	PR	99US-0158369.
PR	13-OCT-1999;	PR	99US-0159293.
PR	13-OCT-1999;	PR	99US-0159294.
PR	13-OCT-1999;	PR	99US-0159295.
PR	14-OCT-1999;	PR	99US-0159329.
PR	14-OCT-1999;	PR	99US-0159330.
PR	14-OCT-1999;	PR	99US-0159331.
PR	14-OCT-1999;	PR	99US-0159637.
PR	14-OCT-1999;	PR	99US-0159638.
PR	18-OCT-1999;	PR	99US-0159584.
PR	21-OCT-1999;	PR	99US-0160741.

PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130892.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 14-MAY-1999; 99US-0134768.
 PR 18-MAY-1999; 99US-0134941.
 PR 19-MAY-1999; 99US-0135124.
 PR 20-MAY-1999; 99US-0135353.
 PR 21-MAY-1999; 99US-0135629.
 PR 24-MAY-1999; 99US-0136392.
 PR 25-MAY-1999; 99US-0136392.
 PR 27-MAY-1999; 99US-0136782.
 PR 28-MAY-1999; 99US-0137222.
 PR 01-JUN-1999; 99US-0137528.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 18-JUN-1999; 99US-0139817.
 PR 18-JUN-1999; 99US-0139817.
 PR 21-JUN-1999; 99US-0139889.
 PR 22-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 23-JUN-1999; 99US-0140695.
 PR 24-JUN-1999; 99US-0140823.
 PR 28-JUN-1999; 99US-0140991.
 PR 29-JUN-1999; 99US-0141287.
 PR 30-JUN-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147260.
 PR 05-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 06-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 23-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151348.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154039.
 PR 16-SEP-1999; 99US-0154779.
 PR 20-SEP-1999; 99US-0155139.
 PR 22-SEP-1999; 99US-0155486.
 PR 23-SEP-1999; 99US-0155659.
 PR 24-SEP-1999; 99US-0156559.
 PR 28-SEP-1999; 99US-0156556.
 PR 29-SEP-1999; 99US-0157117.
 PR 04-OCT-1999; 99US-0157117.

PR	05-OCT-1999;	990S-0157753.
PR	06-OCT-1999;	990S-0157865.
PR	07-OCT-1999;	990S-0158029.
PR	08-OCT-1999;	990S-0158232.
PR	12-OCT-1999;	990S-0158369.
PR	13-OCT-1999;	990S-0159293.
PR	13-OCT-1999;	990S-0159294.
PR	13-OCT-1999;	990S-0159295.
PR	14-OCT-1999;	990S-0159329.
PR	14-OCT-1999;	990S-0159330.
PR	14-OCT-1999;	990S-0159331.
PR	14-OCT-1999;	990S-0159637.
PR	14-OCT-1999;	990S-0159638.
PR	18-OCT-1999;	990S-0159584.
PR	21-OCT-1999;	990S-0160741.
PR	21-OCT-1999;	990S-0160767.
PR	21-OCT-1999;	990S-0160768.
PR	21-OCT-1999;	990S-0160770.
PR	21-OCT-1999;	990S-0160814.
PR	21-OCT-1999;	990S-0160815.
PR	22-OCT-1999;	990S-0160980.
PR	22-OCT-1999;	990S-0160981.
PR	22-OCT-1999;	990S-0160989.
PR	25-OCT-1999;	990S-0161404.
PR	25-OCT-1999;	990S-0161405.
PR	25-OCT-1999;	990S-0161406.
PR	26-OCT-1999;	990S-0161359.
PR	26-OCT-1999;	990S-0161360.
PR	26-OCT-1999;	990S-0161361.
PR	28-OCT-1999;	990S-0161920.
PR	28-OCT-1999;	990S-0161992.
PR	28-OCT-1999;	990S-0161993.
PR	29-OCT-1999;	990S-0162142.

Query Match	32.6%	Score 303.5;	DB 21;	Length 119;
Best Local Similarity	51.3%	Pred. No. 2,2e-24;		
Matches 58;	Conservative 24;	Mismatches 28;	Indels 3;	Gaps 1;
QY	67	QVRSVHLTIKNNOSRRPKSMKSPDG---ISRTRDESIOLKKHLERISLGEVYKLS	ELANT 123	
DB	6	gykashllhllkqgsrkrksawkdpqeklllltttreavaeqklsiredlysgkanfeevatr	65	
QY	124	ESDGSCHDRGCDLGFSSKGQOMQPFEEAFLNLHGEVSNIIETNSGVHILQRT	176	
		: : : :		
DB	66	vsdcssakrrgdlgsfgrgymqkpfceeatyalkygdlsdldvdtsgvnhlkrtr	118	
RESULT 6				
G43433				
ID	G43433	standard;	Protein; 119	AA.
XX	AC			
XX	G43433;			
XX	DT	18-OCT-2000	(first entry)	
XX	DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 54286.	
XX	KW	Protein identification; signal transduction pathway; metabolic pathway;		
XX	KM	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX	RV	termination sequence.		
XX	OS	Arabidopsis thaliana.		
XX	PN	EP1033405-A2.		
XX	PD	06-SEP-2000.		
XX	PF	25-FEB-2000; 2000EP-0301439.		
XX	PR	25-FEB-1999; 99US-0121825.		
XX	PR	05-MAR-1999; 99US-0123348.		
XX	PR	09-MAR-1999; 99US-0123548.		
XX	PR	23-MAR-1999; 99US-0125788.		

PR	25-MAR-1999	9905-01262664
PR	29-MAR-1999	9905-0126785
PR	01-APR-1999	9905-01274652
PR	06-APR-1999	9905-01282334
PR	08-APR-1999	9905-01287134
PR	16-APR-1999	9905-0129845
PR	15-APR-1999	9905-0130077
PR	21-APR-1999	9905-0130449
PR	23-APR-1999	9905-01305510
PR	23-APR-1999	9905-01308691
PR	28-APR-1999	9905-0131448
PR	30-APR-1999	9905-01320448
PR	30-APR-1999	9905-0132407
PR	04-MAY-1999	9905-01324854
PR	05-MAY-1999	9905-01324885
PR	06-MAY-1999	9905-01324876
PR	06-MAY-1999	9905-0132487
PR	07-MAY-1999	9905-01328653
PR	11-MAY-1999	9905-0134256
PR	14-MAY-1999	9905-0134218
PR	14-MAY-1999	9905-0134219
PR	14-MAY-1999	9905-01344211
PR	14-MAY-1999	9905-01343760
PR	18-MAY-1999	9905-01347478
PR	19-MAY-1999	9905-0134041
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135523
PR	24-MAY-1999	9905-01365629
PR	25-MAY-1999	9905-0136621
PR	27-MAY-1999	9905-0136592
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-01375052
PR	07-JUN-1999	9905-0137724
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-01394553
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-01394692
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140685
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140391
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141842
PR	01-JUL-1999	9905-0142154
PR	02-JUL-1999	9905-0142055
PR	06-JUL-1999	9905-0142290
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0142977
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	16-JUL-1999	9905-0144085

9905 - 0137724
9905 - 0138504
9905 - 0138540
9905 - 0138649
9905 - 0138817
9905 - 0138847
9905 - 0139452
9905 - 0139453
9905 - 0139492
9905 - 0139494
9905 - 0139454
9905 - 0139455
9905 - 0139456
9905 - 0139457
9905 - 0139457
9905 - 0139458
9905 - 0139458
9905 - 0139459
9905 - 0139460
9905 - 0139461
9905 - 0139462
9905 - 0139463
9905 - 0139465
9905 - 0139470
9905 - 0139763
9905 - 0139817
9905 - 0139819
9905 - 0139819
9905 - 0140354
9905 - 0140354
9905 - 0140354
9905 - 0140695
9905 - 0140823
9905 - 0140991
9905 - 0141287
9905 - 0141287
9905 - 0141842
9905 - 0142154
9905 - 0142051
9905 - 0142055
9905 - 0142390
9905 - 0142390
9905 - 0142803
9905 - 0142920
9905 - 0142977
9905 - 0143542
9905 - 0143624
9905 - 0144005
9905 - 0144085


```

PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 32.6%; Score 303.5; DB 21; Length 119;
Best Local Similarity 51.3%; Pred. No. 2; 2e-24;
Matches 58; Conservative 24; Mismatches 28; Indels 3; Gaps 1
6 qvkashllikqgstrksaswkdpegklllttreavaeqklsredivsgkanfeevatr 65
QY 67 QVRVSHLLIKNNOSRKPRKSWKSPG--ISRTDESIOIKKHLERILSGEKLSELTANT 123
DB 6 66 vdcssakrvgdlsfgfgrgmqkpleealtakvgydisldvdcsgvhlilkr 118
DE 30-AUG-2000 (first entry)
AC Y90945;
XX Y90945;
DT 30-AUG-2000 (first entry)
XX Cenarchaeum symbiosum open reading frame protein sequence SEQ ID NO:68.
DE Cenarchaeum symbiosum; non-thermophilic; cenarchaeote; physiology;
XX Cenarchaeum symbiosum; non-thermophilic; cenarchaeote; physiology;
KM characterisation; archaea; therapeutic; industrial; laboratory.
XX Cenarchaeum symbiosum.
OS WO200018909-A2.
XX WO200018909-A2.
XX 06-APR-2000.
XX 29-SEP-1999; 99WO-US22752.
XX 29-SEP-1998; 98US-0102294.
XX (DIVE-) DIVERSA CORP.
XX Swanson RV, Feldman RA, Schleper C;
XX WPI; 2000-293148/25.
XX N-PSDB; A55220.
DR

```

XX New nucleic acids and proteins isolated from the non-thermophilic
PT crenarchaeote Cenarchaeum symbiosum, useful in characterizing the
PT physiology of these archaea and in therapeutic, industrial or laboratory
XX techniques -
PS Claim 24; Page 186-187; 210pp; English.
XX
XX A55186 to A55226 and Y90913 to Y90951 represent nucleic acids and
CC proteins isolated from the non-thermophilic crenarchaeote
CC Cenarchaeum symbiosum. The nucleic acids and proteins identified in
CC the present invention are useful in characterizing the physiology of
CC these archaea and can be used in therapeutic, industrial or laboratory
CC techniques. A55227 to A55260 represent promoter sequences from
CC Cenarchaeum symbiosum. A55261 to A55269 represent PCR primers and
CC probes used in examples from the present invention.
XX
SQ Sequence 92 AA;
Query Match 16.8%; Score 157; DB 21; Length 92;
Best Local Similarity 33.9%; Pred. No. 4,1e-09;
Matches 38; Conservative 22; Mismatches 28; Indels 24; Gaps 4;
OY 67 QVRVSHLLIKNNOSRKPSKSPDCISRTDPSIQILKHLERILSGEYKLSLANTES-125
DB 4 kkskshllvkr-----ggqalav-----qerlkage-ktgklakelsl 40
OY 126 DCSHHRGDLGFFSKGQMPPEEAFNLHVGEVSNIIETNSGVHIIQIRFG 177
DB 41 dgsaktrgsglyfgrkmvkrpfedaatrllygvevsepksefyghylikrlg 92
RESULT 8
W89744 W89744 standard; Protein; 320 AA.
AC W89744;
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus protein SEQ ID #5192.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 17..32 /note="these residues represent a line of missing text
FT in the sequence listing in the specification.
FT They are included to maintain the residue
FT numbering given in the specification for this
FT protein sequence"
XX
PN EP786519-A2.
XX
PD 30-JUL-1997.
XX
PE 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Fannon MR, Kunsch CA;
XX
XX Rosen CA;
XX
XX WPI: 1997-374922/35.
XX

PT Polynucleotide(s) and proteins derived from Staphylococcus aureus
PT stored on computer readable medium and used in the production of
PT anti-S.aureus vaccines
XX
PS Claim 23; Page 3187-3188; 3271pp; English.
XX
XX This sequence represents a Staphylococcus aureus protein sequence of the
CC invention. The DNA sequences encoding the S.aureus proteins are recorded
CC on a computer readable medium, preferably selected from a floppy or hard
CC disk, random access memory (RAM), read-only memory (ROM) or CD-ROM.
CC Homology searches using the S.aureus DNA sequences allows putative
CC functions to be assigned so that protein-encoding or regulatory regions
CC of commercial, therapeutic or industrial importance can be obtained.
CC Specifically, sequences which are likely to encode antigens have been
CC identified and these polypeptides can be used in a vaccine composition
CC against S.aureus infection. The polypeptides can also be used in a kit
CC for the immunodetection of S.aureus in a sample. S.aureus is implicated
CC in numerous human diseases, including cellulitis, eyelid infections, food
CC poisoning, osteomyelitis, skin and surgical wound infections, scalded
CC skin syndrome, toxic shock syndrome, etc. Organisms transformed with the
CC DNA sequences can be used for recombinant production of the polypeptides.
CC The new DNA sequences (and their fragments) are useful as primers or
CC probes for isolating homologues of any of the 5191 S.aureus DNA sequences
CC contained on the computer readable medium.
XX
SQ Sequence 320 AA;
Query Match 15.3%; Score 142.5; DB 18; Length 320;
Best Local Similarity 33.3%; Pred. No. 8,5e-07;
Matches 38; Conservative 23; Mismatches 44; Indels 9; Gaps 4;
OY 62 VNEDGOVVRVSHLLIKNNOSRKPSKSPDCISRTDPSIQILKHLERILSGEYKLSLA 121
DB 137 lkeds-k-kashllik-vsksksdexglldkkaekgaeiq-----kvsdpdkfgeia 188
OY 122 NTES-DCSSHHRGDLGFFSKGQMPPEEAFNLHVGEVSNIIETNSGVHIIQ 174
DB 189 kkesmdtgsakkdgelgyvlkgqtdfekalflkdgvevsvksfyghylik 242
RESULT 9
G08943 G08943 standard; Protein; 221 AA.
XX
AC G08943;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 6677.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR

Fri Jun 8 09:43:27 2001

us-09-507-242-2.rag

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0114449.
PR 30-APR-1999; 99US-0120048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132404.
PR 05-MAY-1999; 99US-0122485.
PR 06-MAY-1999; 99US-0132487.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137724.
PR 07-JUN-1999; 99US-0138094.
PR 08-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 10-JUN-1999; 99US-0139119.
PR 14-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139817.
PR 18-JUN-1999; 99US-0139899.
PR 21-JUN-1999; 99US-0140353.
PR 22-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140595.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142380.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143542.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 13-AUG-1999; 99US-0149368.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149422.
PR 18-AUG-1999; 99US-0149422.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0150566.
PR 25-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 31-AUG-1999; 99US-0151930.
PR 01-SEP-1999; 99US-0152363.
PR 07-SEP-1999; 99US-0153070.
PR 10-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154779.
PR 20-SEP-1999; 99US-0155139.
PR 22-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158368.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR	30-APR-1999	9905-01324.07
PR	04-MAY-1999	9905-01324.84
PR	05-MAY-1999	9905-01324.85
PR	06-MAY-1999	9905-01324.86
PR	06-MAY-1999	9905-01324.87
PR	07-MAY-1999	9905-01324.863
PR	11-MAY-1999	9905-01324.65
PR	14-MAY-1999	9905-01324.18
PR	14-MAY-1999	9905-01342.19
PR	14-MAY-1999	9905-01342.19
PR	14-MAY-1999	9905-01344.21
PR	14-MAY-1999	9905-01343.70
PR	18-MAY-1999	9905-01347.68
PR	19-MAY-1999	9905-01334.41
PR	20-MAY-1999	9905-01351.24
PR	21-MAY-1999	9905-01353.53
PR	24-MAY-1999	9905-01356.29
PR	25-MAY-1999	9905-01360.21
PR	27-MAY-1999	9905-01363.92
PR	01-JUN-1999	9905-01367.82
PR	03-JUN-1999	9905-01372.22
PR	03-JUN-1999	9905-01375.28
PR	04-JUN-1999	9905-01375.02
PR	07-JUN-1999	9905-01377.24
PR	08-JUN-1999	9905-01380.94
PR	10-JUN-1999	9905-01385.40
PR	10-JUN-1999	9905-01388.47
PR	14-JUN-1999	9905-01391.19
PR	16-JUN-1999	9905-01394.52
PR	16-JUN-1999	9905-01394.53
PR	17-JUN-1999	9905-01394.92
PR	18-JUN-1999	9905-01394.54
PR	18-JUN-1999	9905-01394.55
PR	18-JUN-1999	9905-01394.56
PR	18-JUN-1999	9905-01394.57
PR	18-JUN-1999	9905-01394.58
PR	18-JUN-1999	9905-01394.59
PR	18-JUN-1999	9905-01396.11
PR	18-JUN-1999	9905-01396.61
PR	18-JUN-1999	9905-01394.62
PR	18-JUN-1999	9905-01394.63
PR	18-JUN-1999	9905-01397.50
PR	21-JUN-1999	9905-01398.17
PR	22-JUN-1999	9905-01400.91
PR	23-JUN-1999	9905-01400.93
PR	23-JUN-1999	9905-01400.94
PR	24-JUN-1999	9905-01405.54
PR	28-JUN-1999	9905-01408.23
PR	28-JUN-1999	9905-01408.23
PR	30-JUN-1999	9905-01412.87
PR	01-JUL-1999	9905-01414.14
PR	01-JUL-1999	9905-01412.54
PR	02-JUL-1999	9905-01420.50
PR	06-JUL-1999	9905-01423.95
PR	08-JUL-1999	9905-01428.20
PR	09-JUL-1999	9905-01428.20
PR	12-JUL-1999	9905-01429.77
PR	13-JUL-1999	9905-01433.42
PR	14-JUL-1999	9905-01435.24
PR	15-JUL-1999	9905-01440.05
PR	16-JUL-1999	9905-01440.05
PR	16-JUL-1999	9905-01440.86
PR	19-JUL-1999	9905-01443.25
PR	19-JUL-1999	9905-01443.31
PR	19-JUL-1999	9905-01443.32
PR	19-JUL-1999	9905-01443.32
PR	19-JUL-1999	9905-01443.33
PR	19-JUL-1999	9905-01443.34
PR	20-JUL-1999	9905-01446.32
PR	20-JUL-1999	9905-01446.32
PR	21-JUL-1999	9905-01446.34
PR	21-JUL-1999	9905-01450.86

PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145085
PR	22-JUL-1999	9905-0145089
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145145
PR	23-JUL-1999	9905-0145148
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513
PR	27-JUL-1999	9905-0145519
PR	28-JUL-1999	9905-0145551
PR	02-AUG-1999	9905-0146386
PR	02-AUG-1999	9905-0146389
PR	03-AUG-1999	9905-0147038
PR	04-AUG-1999	9905-0147204
PR	05-AUG-1999	9905-0147192
PR	05-AUG-1999	9905-0147260
PR	06-AUG-1999	9905-0147703
PR	06-AUG-1999	9905-0147416
PR	09-AUG-1999	9905-0147493
PR	09-AUG-1999	9905-0147735
PR	10-AUG-1999	9905-0148171
PR	11-AUG-1999	9905-0148319
PR	12-AUG-1999	9905-0148341
PR	13-AUG-1999	9905-0148565
PR	13-AUG-1999	9905-0148684
PR	15-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	20-AUG-1999	9905-0149722
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149929
PR	23-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	25-AUG-1999	9905-0150564
PR	26-AUG-1999	9905-0150684
PR	27-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151080
PR	30-AUG-1999	9905-0151133
PR	31-AUG-1999	9905-0151338
PR	01-SEP-1999	9905-0151330
PR	01-SEP-1999	9905-0151263
PR	10-SEP-1999	9905-0153750
PR	13-SEP-1999	9905-0153758
PR	13-SEP-1999	9905-0154018
PR	15-SEP-1999	9905-0154779
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155131
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155659
PR	26-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159320
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741

PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145192;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145224;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	02-AUG-1999;	9905-0147038;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148311;
PR	12-AUG-1999;	9905-0148355;
PR	13-AUG-1999;	9905-0148561;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-01493368;
PR	17-AUG-1999;	9905-01493175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	23-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149930;
PR	23-AUG-1999;	9905-0149930;
PR	26-AUG-1999;	9905-0150566;
PR	27-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	30-AUG-1999;	9905-0151083;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-0151930;
PR	07-SEP-1999;	9905-0155363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154018;
PR	20-SEP-1999;	9905-0154179;
PR	22-SEP-1999;	9905-0155131;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155569;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156566;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;
PR	07-OCT-1999;	9905-0158029;
PR	08-OCT-1999;	9905-0158232;
PR	13-OCT-1999;	9905-0158363;
PR	13-OCT-1999;	9905-0158293;
PR	13-OCT-1999;	9905-0159294;
PR	14-OCT-1999;	9905-0159329;
PR	14-OCT-1999;	9905-0159330;
PR	14-OCT-1999;	9905-0159331;
PR	14-OCT-1999;	9905-0159637;
PR	14-OCT-1999;	9905-0159638;
PR	18-OCT-1999;	9905-0159584;
PR	21-OCT-1999;	9905-0160741;
PR	21-OCT-1999;	9905-0160767;
PR	21-OCT-1999;	9905-0160770;

[illegible]

PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0144218.
 PR 14-MAY-1999; 99US-0144219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139464.
 PR 18-JUN-1999; 99US-0139465.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140823.
 PR 30-JUN-1999; 99US-0140823.
 PR 01-JUL-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144632.
 PR 21-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145919.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146388.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 10-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 26-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 01-SEP-1999; 99US-0152363.
 PR 01-SEP-1999; 99US-0153070.
 PR 10-SEP-1999; 99US-0153758.
 PR 13-SEP-1999; 99US-0154018.
 PR 15-SEP-1999; 99US-0154039.
 PR 16-SEP-1999; 99US-0154779.
 PR 20-SEP-1999; 99US-0155139.
 PR 22-SEP-1999; 99US-0155486.
 PR 23-SEP-1999; 99US-0155659.
 PR 24-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159283.
 PR 13-OCT-1999; 99US-0159285.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159641.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.

PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 13.7%; Score 127.5; DB 21; Length 169;
 Best Local Similarity 30.1%; Pred. No. 1.3e-05;
 Matches 34; Conservative 21; Mismatches 41; Indels 17; Gaps 2;

Oy 82 KPKSMKSPDGIS-----RTRDESIOLKKHLERLISGEVKL-----SELANTE 124
 Db 54 kxgaqkaadgigtctyvarhvlcekgklineaykklqdgwlsngdkvypaefaklaeay 113
 Oy 125 SDCSSHDRGDLGFFFSKQOMOPPEEAFLNLHVGESVNIETNSGVHILQRTG 177
 Db 114 secpsgkkggdgwfprxmagpfgdvaintrpyvtstapfkstngyhllseg 166

RESULT 14
 W89814
 ID W89814 standard; Protein; 204 AA.
 AC W89814;

DT 18-FEB-1999 (first entry)
 DE Protein encoded by clone D7.

KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
 KW Peptic ulcer; gastric adenocarcinoma; gastric lymphoma.

OS Helicobacter pylori.

PN W09849314-A2.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-US08487.

PR 14-OCT-1997; 97US-0061958.

PR 25-APR-1997; 97US-0045107.

PA (GENE-) GENELABS TECHNOLOGIES INC.

PI Chow TP, Fry KE, Lim MY, McAttee CP;

DR WPI: 1999-009433/01.

N-PSDB; V90546.

PT New Helicobacter pylori antigens and related nucleic acid sequences
 PT - useful in serological diagnosis and protective vaccines, providing
 PT long-lasting immune response

PS Claim 15; Page 97-98; 402pp; English.

CC The present sequence represents a Helicobacter pylori antigenic protein
 CC that is characterised by immunoreactivity with H. pylori-positive
 CC antisera. The proteins are highly immunogenic and induce a long-lasting
 CC immune response that persists even after antimicrobial treatment. In
 CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
 CC highly sensitive and specific. The specification also describes 69
 CC previously unrecognised immunogenic cluster families. H. pylori antigens
 CC are used to detect H. pylori-specific antibodies, for diagnosing
 CC infection or to confirm eradication of infection, and in vaccines to
 CC protect against H. pylori infection and related diseases (gastritis,

CC peptic ulcer, gastric adenocarcinoma/lymphoma).
 XX SQ Sequence 204 AA;

Query Match 13.0%; Score 121.5; DB 20; Length 204;
 Best Local Similarity 29.2%; Pred. No. 7.2e-05;
 Matches 40; Conservative 23; Mismatches 45; Indels 29; Gaps 6;

Oy 42 DKEVLNAVIAEFKNGYKPLVNEQVRSVSHLLIKRNSRKPMSKSPDGISRTRDESIQ 101
 Db 65 ekemqdfyna-----nkqqlfvtqgeahar--hlvk-----tedekr 100
 Oy 102 ILKK-HLEERLISGEVKSELANTES---DCSSHDRGDLGFFFSKQOMOPPEEAFLNLHV 157
 Db 101 llseidkqpkakeekfielelnrdtldpnskaqngsgllgkqkqmpdskaafalts 160

Oy 158 GEVSNL-IEVNSGVHIL 173
 Db 161 gdytktpvktefgyhll 177

RESULT 15
 W55450
 ID W55450 standard; Protein; 299 AA.
 AC W55450;

DT 24-JUN-1998 (first entry)

DE H. pylori ORF 02ae11612_22477267_f2_27 cell envelope OMP.

KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
 KW Identification; binding compound; bacteria; life cycle; activator;
 KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis.

OS Helicobacter pylori.

PN W09737044-A1.

PD 09-OCT-1997.

PF 27-MAR-1997; 97WO-US05223.

PR 06-DEC-1996; 96US-0761318.

PR 29-MAR-1996; 96US-0625811.

PR 02-APR-1996; 96US-0758731.

PR 25-OCT-1996; 96US-0736905.

PR 28-OCT-1996; 96US-0738859.

PA (ASTR) ASTRA AB.

PI Alm RA, Smith D;

DR WPI: 1997-503122/46.

N-PSDB; V24859.

PT Helicobacter pylori nucleic acid sequences and encoded
 PT polypeptide(s) - useful in vaccines to treat or prevent H. pylori
 PT infection and for diagnosis of H. pylori infection

PS Claims 14,80; Page 657; 1145pp; English.

CC This sequence is a H. pylori cell envelope outer membrane protein
 CC (OMP) having no terminal phe residue.
 CC The protein may be used in a vaccine to prevent or treat H. pylori
 CC infection or to identify H. pylori polypeptide binding compounds, the
 CC useful as potential H. pylori life cycle activators or inhibitors.
 CC The DNA and probes derived from it may be used for the identification of
 CC H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
 CC acid sequences complementary to the DNA act as antisense sequences and
 CC can be used to prevent the translation of H. pylori mRNA. Antibodies
 CC against the protein can be used in immunoassays to evaluate the abundance

...

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2001, 09:37:50 ; Search time 14.18 Seconds
(Without alignments)
857,825 Million cell updates/sec

Title: US-09-507-242-2

Perfect score: 932
Sequence: 1 MASTSTGLPPNMTIRVSRSH.....GEVSNIIETNSGVHLIORTG 177

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR-67:*

1: p1r1:*

2: p1r2:*

3: p1r3:*

4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	409	43.9	175	T41093	peptidyl-prolyl ci
2	369.5	39.6	163	S68520	pini protein - human
3	367	39.4	170	S52764	ESS1 protein - yeast
4	361.5	38.8	165	JC7136	peptidyl-prolyl iso
5	325	34.9	166	T08426	pini protein homol
6	170.5	18.3	292	S15269	post-translocation
7	166.5	17.9	333	A83797	protein secretion
8	142	15.2	221	T52622	probable peptidyl
9	124	13.3	621	A83412	peptidyl-prolyl ci
10	123	13.2	273	S52412	major antigenic pe
11	122	13.1	347	A81794	probable rotamase
12	122	13.1	348	B81216	peptidyl-prolyl ci
13	121.5	13.0	299	G64541	cell binding facto
14	121.5	13.0	299	B71967	probable peptidyl-
15	121	13.0	93	S48658	peptidyl-prolyl iso
16	119.5	12.8	92	A82396	peptidyl-prolyl ci
17	118.5	12.7	655	E82712	peptidyl-prolyl ci
18	117.5	12.6	282	E71662	protein export pro
19	116.5	12.5	336	D70113	basic membrane pro
20	116.5	12.5	619	E82141	peptidyl-prolyl ci
21	110.5	11.9	594	B64018	hypothetical prote
22	110.5	11.8	430	B83572	peptidyl-prolyl ci
23	109	11.7	93	C83123	peptidyl-prolyl ci
24	108	11.6	431	D82323	survival protein S
25	106	11.4	272	C83162	probable peptidyl-
26	105	11.3	292	F32055	nifm protein - Azo
27	104.5	11.2	92	F83395	peptidyl-prolyl ci
28	103	11.1	293	H43706	nifm protein - Azo
29	97.5	10.5	297	S66102	protein secretion

30	95	10.2	126	2	T31601	hypothetical prote
31	89	9.5	428	2	B64726	probable peptidylp
32	88.5	9.5	314	2	B83559	probable peptidyl-
33	87.5	9.4	464	2	C82756	peptidyl-prolyl ci
34	85.5	9.2	879	2	E71811	probable iron-regu
35	85	9.1	1713	2	A55347	adhesive ligand ep
36	84.5	9.1	583	2	S30014	hypothetical prote
37	84.5	9.1	709	2	T16584	hypothetical prote
38	84.5	9.1	877	2	H64708	iron-regulated out
39	84	9.0	150	1	G69873	ynuv-related prote
40	83.5	9.0	911	2	S51441	hypothetical prote
41	81.5	8.7	241	2	T25886	hypothetical prote
42	81	8.7	483	2	T19720	hypothetical prote
43	80.5	8.6	302	2	F64350	ribokinase (EC 2.7
44	80.5	8.6	468	2	T21331	hypothetical prote
45	80.5	8.6	1013	2	T46422	hypothetical prote

ALIGNMENTS

RESULT 1
T41093
peptidyl-prolyl cis-trans isomerase - fission yeast (*Schizosaccharomyces pombe*)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000
C:Accession: T41093
R:Rumelle, B.; Goffeau, A.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, September 1998
A:Reference number: 221964
A:Accession: T41093
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-175 <PDB>
A:Cross-references: EMBL:AL031535; PIDN:CAA20742.1; GSPDB:GN00068; SPDB:SPCC16C4.03
A:Experimental source: EMBL: strain 972h-; cosmid c16C4
A:Genetics:
A:Gene: SPDB:SPCC16C4.03
A:Map position: 3
A:introns: 134/2
A:Superfamily: yeast ESS1 protein; WW repeat homology
F:4-42/Domain: WW repeat homology <WWR>

Query Match 43.9% Score 409; DB 2; Length 175;
Best local Similarity 46.0%; Pred. No. 4.6e-28;
Matches 80; Conservative 37; Mismatches 53; Indels 4; Gaps 4;
QY 4 TSTGLPPNMTIRVSRSHNKEFLNOSTNESSWDPPYGTDEVLNAYIAKFNNGYKPL-V 62
DB 2 SNTGLKRPVAVKISRSRNPFFENTHESLWEPAAIDMALKKFIINELQESVTPTEA 61
QY 63 NEDQVAVSHLTKNNOSRKPMSKSPDGISTRROESIOILKKHLERIL-SGEVNLSELA 121
DB 62 SNPKIRASHLTKHRSRSPSSWKE-EHITNSKEEA-RKLAHEHQLKSSVSVMHDLA 119
QY 122 NTESDCSHDRGDLGFGSKGQMQPPEEAAPNLHAGEVSNIIETNSGVHLIORTG 175
DB 120 MKESDCSARRGELGFGEDMKPFEDAALFKGEISGVETSSGPHIORT 173
RESULT 2
S68520
Pini protein - human
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1996 #sequence_revision 27-Feb-1997 #text_change 29-Sep-1999
C:Accession: S68520
R:Lin, K.P.; Hanes, S.D.; Hunter, T.
Nature 380, 544-547, 1996
A:Title: A human peptidyl-prolyl isomerase essential for regulation of mitosis.
A:Reference number: S68520; MUID:96195064
A:Accession: S68520
A:Status: preliminary; nucleic acid sequence not shown

```

A:Molecule type: mRNA
A:Residues: 1-163 <LNU>
A:Cross-references: EMBL:U049070; NID:g1332709; PIDN:AAC50492.1; PID:g1332710
C:Genetics:
A:Gene: GDB:PIN1; docd
A:Cross-references: GDB:521831
C:Superfamily: yeast ESS1 protein; MW repeat homology
F:5-43/Domain: MW repeat homology <MWI>

Query Match          39.6%; Score 369.5; DB 2; Length 163;
Best Local Similarity 43.2%; Pred. No. 1e-24;
Matches 73; Conservative 34; Mismatches 49; Indels 13; Gaps 2;

QY      8 LPPMNTIVSSHNKEYELYNOSTNESSMDPYGTDEVLNAVIAKFNKGKPLVNEGQ 67
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |
Db       7 LPFGMEKRMSRSSGGRVYYFNHTTNASOWERPSP-----SSSGGNNGGEPAR 54
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |

QY     68 VRVSHLLIKNNOSRPRKRWKSPDGISRTRODSIQLKHLERILSGEYKLSLATNESDC 127
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |
Db     55 VRCSHLLVKHSOSRRPSSWRO-EKITRKKEALELINCYIQRKIGSEDFESLASQFSDC 113
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |

QY    128 SSHDGGGLGFSSGOMOPPEEAFAFLNHGVSAFIIFTSNVGHLORT 176
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |
Db   114 SSARAGRLGAFSRGQMOKPFEDNASFALRTGEMSGPVFTDGIHILLRT 162
        |||::|||::| | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 3

SSSI protein - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein J1452; protein YJR017C; Pfl1 protein

C:Species: *Saccharomyces cerevisiae*

C:Date: 19-May-1995 #sequence,revision 01-Sep-1995 #text_change 29-Sep-1999

C:Accession: S52764; S55207; S57032; S07867; S65929

R:Hani, J.; Stumpf, G.; Domdey, H.

R:Submitted to the EMBL Data Library, March 1995

R:Description: Pfl1 encodes an essential protein in *Saccharomyces cerevisiae*, which shows

A:Reference number: S52762

A:Accession: S52764

A:Molecule type: DNA

A:Residues: 1-170 <HAN>

A:Cross-references: EMBL:X65972; NID:g758283; PIDN:CAA59961.1; PID:g758286

R:de Haan, M.; Smits, P.H.M.; Grivell, L.A.

R:Submitted to the EMBL Data Library, May 1995

A:Reference number: S55183

A:Accession: S55207

A:Molecule type: DNA

A:Residues: 'MTIVVLIPTNRLRHNDAP', 1-170 <DEH>

A:Cross-references: EMBL:X87611; NID:g854567; PIDN:CAA60941.1; PID:g854592

R:de Haan, M.; Grivell, L.A.; Smits, P.H.M.

R:Submitted to the Protein Sequence Database, September 1995

A:Reference number: S56771

A:Accession: S57032

A:Molecule type: DNA

A:Residues: 'MTIVVLIPTNRLRHNDAP', 1-170 <ZAG>

A:Cross-references: EMBL:Z49517; NID:g1015651; PIDN:CAA89541.1; PID:g1015652; MIPS:YJR01

R:Hanes, S.D.; Shank, P.R.; Bostian, K.A.

R:Yeast 5, 55-72, 1989

A:Title: Sequence and mutational analysis of SSSI, a gene essential for growth in *Sacche*

A:Reference number: S07867; MUID:89189095

A:Accession: S07867

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 'MTIVVLIPTNRLRHNDAP', 1-170 <S>

A:Experimental source: strain DBY864

R:Hani, J.; Stumpf, G.; Domdey, H.

R:FEBS Lett. 365, 198-202, 1995

A:Title: Pfl1 encodes an essential protein in *Saccharomyces cerevisiae*, which shows stric

A:Reference number: S65927; MUID:95300974

A:Accession: S65929

A:Molecule type: DNA

A:Residues: 1-170 <HAN>

A:Cross-references: EMBL:X65972; NID:g758283; PIDN:CAA59961.1; PID:g758286

C:Genetics:
A:Gene: SGD:ESS1; PTF1
A:Cross-references: SGD:S0003778; MIPS:YJR017c
A:Map position: 10R
C:Superfamily: Yeast ESS1 protein; WW repeat homology
F:9-47/domain: WW repeat homology <WW1>

```

Query March 39.4% Score 367; DB 2; Length 170;
Best Local Similarity 43.3%; Pred. No. 1.8e-24;
Matches 77; Conservative 35; Mismatches 50; Indels 16; Gaps 4;

QY 3 STSTGLPNNMTIRVSRSHNKEVFLNOSTNESSMPPYGTDEVLNATYAKERNNGYKPLV 62
   :: :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 6 ASRTGLPPIWIVRRYSKSKAKREYFNFPEIKHSQWMEPECTNQDL-----HKILR 54

QY 63 NEDGQVRVSHLLIKNNQSRKPKSMKSPGICSTRDESIQIILAKHLERLSEVKLSF--- 119
   :: :|| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 55 DHPVVRCLHLILIKRDSRRPASRHS -NITISKQDAADKLTILTR-LDDDSKTNFSPEA 112

QY 120 LANFSPDSSHDHGGDLGFSGQGMQPFEEAALFVHGEVSNITETSGVHIIORTG 177
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 113 LAKESSDCSSIKRGDDLFGFGEVQMPSEFEDAAALQKAGEVSDIYESSGVHYIKRVG 170

```

```

RESULT      4
JC7136
peptidylprolyl isomerase (EC 5.2.1.8) - mouse
N:Alternate names: parvulins, Pnl1 protein
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 18-Aug-2000
C:Accession: JC7136
R:Fujimori, F.; Takahashi, K.; Uchida, C.; Uchida, T.
Biochem. Biophys. Res. Commun. 265, 658-663, 1999
A:Title: Mice lacking Pnl1 develop normally, but are defective in entering cell cycle
A:Reference number: JC7136; MUID:20070807
A:Accession: JC7136
A:Reference type: mRNA
A:Molecule type: mRNA
A:Residues: 1-165 <FUJ>
A:Cross-references: DDBJ:AB009691; NID:g6468199; PIDD:BAA87037.1; PTD:g6468200
C:Comment: This protein is a member of peptidyl prolyl cis/trans isomerase family. It
tes in oligopeptides, controls G2/M transition of HeLa cells, regulates entry into G1
C:Genetics:
A:Gene: Pnl1
C:Superfamily: yeast ESS1 protein; WW repeat homology
C:Keywords: cis-trans-isomerase; mitosis
P:5-43/Domain: WW repeat homology <MMR>

```

```

Query Match          38.8%  Score 361.5;  DB 2;  Length 165;
Best Local Similarity 42.68;  Pred. No. 5e-24;
Matches 72;  Conservative 35;  Mismatches 51;  Indels 11;  Gaps 3;

QY      8  LPPMWTITVSSSHNKEVFLNOSTNESSMDPPYGTDKEVLNAYIAKFKNNGKPYLVNEDGQ 67
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       7  LPPGEMKRMSSSSGKRVYYFNHITNASSQWERPSSG-----GSTVGSSSKNGQ-----GEPAK 56

QY      68  VRSVSLILKNNQSRKPKKWKSPDGSRRRDSIQILKHLERLISGEVKSSELANTESDC 127
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       57  VRCSHLIAKHSSQSRPSSWRO-EKITSRKEALELILNELYIOKIKSGEDFESLASQFSDC 115
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY      128  SSHDGGDGLGFFGSKGOMQPFEEAALFNHVEVSNIIETNSGVHILORT 176
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db       116  SSAKARGDLGPFSSKGQMKPFEDASFLAKRTDEMGSQPVTTDSGHILIRT 164

RESULT      5
T084Z6
Pini1 protein homolog dodo - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 17-Nov-2000
C:Accession: T084Z6
C:Matesszka, R.; de Couet, H.G.; Miklos, G.L.

```

Accession: J40002 translated from GB/EMBL/DBJ
A:Status: preliminary
A:Accession type: DNA
A:Residues: MKRSR, 4, LL, 7-8, VIIS, 13, VESTVITQ, 22,181-195, T, 197-210, LE, 279-280, A
A:Cross-References: GB:M22903; NID:9143687; PIDN:AAA22825.1; PID:9143688
R:Runt, F.; Ogasawara, M.; Moszer, I.; Albertini, A.M.; Allioni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Broutlet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chid
A.: Enlich, S.D.; Emerson, P.T.; Erlton, K.D.; Erlington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallazzi, A.; Gallen
leeh, J.; Harwood, C.R.; Hnault, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulo, M.F.
Koetler, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kuita, A.; Liu, H.; Masuda, S.; Meusel
A:Authors: Iabner, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Meusel
M.; Ogawa, K.; Ojima, A.; Oudega, B.; Park, S.H.; Paro, V.; Pohl, T.M.; Portetell
Rieger, M.; Rivolta, C.; Roche, B.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A:Authors: Schleicher, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Serovica, A.;
Knecht, M.; Tamkoshi, A.; Tanaka, T.; Terstap, P.; Tognoni, A.; Tostko, V.; Uchiyama, K.
T.; Winters, P.; Wipelt, A.; Yamamoto, H.; Yaman, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A:Authors: Yoshikawa, H.F.; Zundstein, E.; Yoshikawa, H.; Zandchin, A.

```

Query Match          17.9%; Score 166.5; DB 2; Length 333;
Best Local Similarity 30.1%; Pred. No. 6,4e-07;
Matches 50; Conservative 25; Mismatches 40; Indels 51; Gaps 7;

QY      21 NKEEPLFNGSNSSWMPGPGTDR-----EVLNAYIAKFKNKGKPLVNEGG--OV 68
        |||::||:         |:-:||:         :||:         :||:
Db       114 NKEEFNVEYL-----PVLYLEKLAEGVEITDEEKQAFYD-----NRSLIEV 158
        |||::||:         :||:         :||:         :||:
QY      69 RVSHLLIKNNQSKPKPSWKSPDGISIRDSIOILKKHLERLISGEVKSELANTES--DC 127
        |||::||:         :||:         :||:         :||:
Db       159 EASHIIVDEET-----AEVLDRLLEAG--DEAELLASEVSYPD 195
        |||::||:         :||:         :||:         :||:
QY     128 SSMHROGDIGFFSGKQMPPFEFAFNLIHAVGVSNIIETNSGVHIL 173
        |||::||:         :||:         :||:         :||:
Db       196 SAENANGDLGFEGKDMVPPEFEFAFMNIEDSEVPYSTGYHIIT 241
        |||::||:         :||:         :||:         :||:

RESULT      8
T52622
Probable peptidylprolyl isomerase (EC 5.2.1.8) [imported] - Arabidopsis thaliana
N_Alternate names: Probable peptidyl prolyl cis-trans isomerase
C_Species: Arabidopsis thaliana (mouse-ear cress)
C_Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C_Accession: T52622
```

R:Rabbiyichuk, E.; Fungtung, M.; Van Montagu, M.; Inze, D.; Kushnir, S.
Proc. Natl. Acad. Sci. U.S.A. 94, 12722-12727, 1999
A:Title: Efficient gene tapping in Arabidopsis thaliana using a gene trap approach.
A:Reference number: 226135
A:Accession: T52622
A:Status: preliminary: translated from GB/EMBL/DDBT
A:Molecule type: mRNA
A:Residues: 1-221 <BAB>
A:Cross-references: EMBL:Z66095; PIDD:CA06699.1
A:Experimental source: cultivar Columbia
A:Keywords: cis-trans-Isomerase

R:Rahfeld, J.; Ruecknagel, K.P.; Schelbert, B.; Ludwig, B.; Hacker, J.; Mann, K.; Fische
 A:Description: Confirmation of the existence of a third family among peptidyl-prolyl cis
 A:Reference number: S45525
 A:Accession: S45525
 A:Molecule type: protein
 A:Residues: 2-93 <RAH1>
 A:Experimental source: strain K-12
 R:Rahfeld, J.U.; Schlerhorn, A.; Mann, K.; Fischer, G.
 FEBS Lett. 343, 65-69, 1994
 A>Title: A novel peptidyl-prolyl cis/trans isomerase from Escherichia coli.
 A:Reference number: S43654; MUID:94215709
 A:Accession: S43654
 A:Molecule type: protein
 A:Residues: 2-22 <RAH2>
 A:Experimental source: strain K-12
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC
 .A.; Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of Escherichia coli K-12.
 A:Reference number: A64720; MUID:97426617
 A:Accession: B65181
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-93 <BLAT>
 A:Cross-references: GB:AE00454; GB:U00096; MID:92367278; PIDN:AAC76780.1; PID:91790211;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: PplC
 C:Keywords: cis-trans-isomerase; cyclosporin A binding

Query Match 13.0%; Score 121; DB 2; Length 93;
 Best Local Similarity 37.3%; Pred. No. 0.00096;
 Matches 31; Conservative 13; Mismatches 31; Indels 8; Gaps 2;
 QY 99 SIQILKKH-----LERLSGEVKLSLEANTESDCSSHDRGDLGFFSKGOMOPPEEA 151
 Db 7 ALHILVKEEKALDLLEQIKNG-ADFGKLAKRHSICPSGRGDLGEFRGOMVPAFDKV 65
 QY 152 AFNLHVGESVNIETNSGVHIO 174
 Db 66 VESCPVLEPTGPLHTQFGYHITK 88

Search completed: June 8, 2001, 09:38:11
 Job time: 21 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 8, 2001, 09:37:55 ; Search time 9.54 Seconds
(without alignments)
635.558 Million cell updates/sec

Title: US-09-507-242-2

Perfect score: 932

Sequence: 1 MASTSTGLPMTIRVSRSH.....GEVSNITETSGVHILQRTG 177

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	369.5	39.6	163	1	PIN1_HUMAN
2	367	39.4	190	1	ESS1_YEAST
3	335	34.9	166	1	DOD_DROME
4	170.5	18.3	292	1	PRSA_BACSU
5	123	13.2	273	1	CBE2_CAMJE
6	121.5	13.0	299	1	Y175_HELPU
7	121.5	13.0	299	1	Y175_HELPU
8	121	13.0	299	1	Y175_HELPU
9	114	12.2	100	1	PIN1_HUMAN
10	110.5	11.9	622	1	PIN1_HUMAN
11	105	11.3	292	1	NIFM_AZOVI
12	103	11.1	293	1	NIFM_AZOVI
13	97.5	10.5	297	1	YACD_BACSU
14	97	10.4	577	1	BAG3_MOUSE
15	89	9.5	428	1	SURA_ECOLI
16	85	9.1	113	1	LMA3_HUMAN
17	84.5	9.1	583	1	PR40_YEAST
18	81.5	8.7	1914	1	KMLS_HUMAN
19	80.5	8.6	302	1	Y406_METRA
20	80.5	8.6	421	1	GAS7_MOUSE
21	80.5	8.6	746	1	EXT1_HUMAN
22	80.5	8.6	746	1	EXT1_HUMAN
23	80	8.6	918	1	YMBJ_CAEEL
24	79.5	8.5	957	1	NED4_MOUSE
25	79	8.5	430	1	SURA_BUCAT
26	79	8.5	1814	1	TSC2_MOUSE
27	78.5	8.4	426	1	NIFM_KLEPN
28	78.5	8.4	422	1	GAS7_RAT
29	78.5	8.4	1664	1	RPA1_YEAST
30	78.5	8.4	2259	1	CCAE_RABIT
31	78	8.4	256	1	M3R_DICDI
32	78	8.4	313	1	SURA_HAEIN
33	78	8.4	387	1	UPF3_YEAST

34	77.5	8.3	724	1	Y061_CAEEL
35	77	8.3	259	1	KLK3_RAT
36	77	8.3	299	1	PRM1_LACPA
37	77	8.3	412	1	GAS7_HUMAN
38	76	8.2	549	1	PP01_YEAST
39	75.5	8.1	182	1	HSLV_BORBU
40	75.5	8.1	361	1	WDS_DROME
41	75	8.0	259	1	KLK3_RAT
42	75	8.0	955	1	MMU3_MYCLE
43	75	8.0	1022	1	PIB4_HUMAN
44	74.5	8.0	228	1	YDZ1_SCHPO
45	74.5	8.0	652	1	MX1_RAT

ALIGNMENTS

RESULT 1
ID PIN1_HUMAN STANDARD; . PRT; 163 AA.
AC 013526;
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE NIMA-INTERACTING 1 (EC 5.2.1.8).
GN PIN1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TAXID:9606;
RN 1
RP SEQUENCE FROM N.A.
RA MEDLINE:96195064; PubMed-8606777;
RT Lu K.P., Hanes S.D., Hunter T.;
RT "A human peptidyl-prolyl isomerase essential for regulation of
mitosis.";
RL Nature 380:544-547(1996)
RN 1
RP X-RAY CRYSTALLOGRAPHY (1.35 ANGSTROMS).
RX MEDLINE:97344079; PubMed-9200606;
RA Ranganathan R., Lu K.P., Hunter T., Noel J.P.;
RT "Structural and functional analysis of the mitotic rotamase Pin1
suggests substrate recognition is phosphorylation dependent.";
RL Cell 89:875-886(1997).
CC -1- FUNCTION: ESSENTIAL PEPIASE THAT REGULATES MITOSIS PRESUMABLY BY
INTERACTING WITH NIMA AND ATTENUATING ITS MITOSIS-PROMOTING
ACTIVITY.
CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- SIMILARITY: BELONGS TO THE PIC/PAVULIN FAMILY OF ROTAMASES.
CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement. (See <http://www.isb.ch/announce/>
or send an email to license@isb.slb.ch.)
CC EMBL: U49070; AAC50492.1; -
CC PDB: 1PIN; 25-NOV-98.
CC MIM: 601052; -
CC InterPro: IPR000297; -
CC InterPro: IPR001202; -
CC InterPro: IPR002349; -
CC Pfam: PF00639; Rotamase; 1.
CC Pfam: PF00397; WW; 1.
CC PRINTS: PR00403; WMDOMAIN.
CC PROSITE: PS01096; PIC_PPIASE_1; 1.
CC PROSITE: PS0198; PIC_PPIASE_2; 1.
CC PROSITE: PS01159; WW_DOMAIN_1; 1.
CC PROSITE: PS01159; WW_DOMAIN_1; 1.

DR PROSITE: PS50020; WW_DOMAIN_2; 1.
 KW Isomerase; Rotamase; Nuclear protein; Cell cycle; 3D-structure.
 FT DOMAIN 5 39 MW.
 FT DOMAIN 52 163 PIC-LIKE.
 SQ SEQUENCE 163 AA: 18243 MW: 35391AF40B7D1E13 CXC64;

Query Match 39.6%; Score 369.5; DB 1; Length 163;
 Best Local Similarity 43.2%; Pred. No. 1.6e-25;
 Matches 73; Conservative 34; Mismatches 49; Indels 13; Gaps 2;

OY 8 LPPMTFIVSRSHNKEYFLNOSTNESSWDPYGTDEKVLNAYIAFKNGKPYLVNEDGQ 67
 DB 7 LPPGWEKRMSSRGHYVFNNHTNASOMERPSCN-----SSSGKNGCGEPAR 54
 OY 68 VRVSHLLKNNQSRKPKSKSPDGISRTDESIQILKKHLERILSGEYKSLATNESIC 127
 DB 55 VCSHLVYKHSQSRPSSMRQ-EKITRKEALELLENGYIOIKSGEEDFESLASQFSDC 113
 OY 128 SSHDRGDLGFESKGMOPPEEAAFNLVGEVSNIIETNSGVHIIORT 176
 DB 114 SSARKRGDLGAFSGRGMQKPFEDASFLRTGEMSGPVFTDSGIIHLRT 162

RESULT 2
 ESSL_YEAST 2
 ID ESSL_YEAST STANDARD: PRT; 190 AA.

AC P22696;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE ESSL PROTEIN (PROCESSING/TERMINATION FACTOR 1).
 OS ESSL OR PTF1 OR YJ017C OR J1452.
 GN Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OK NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89189095; PubMed=2648698;
 RA Hanes S.D., Shank P.R., Bostian K.A.;
 RT "Sequence and mutational analysis of ESSL, a gene essential for
 RL growth in Saccharomyces cerevisiae.";
 Yeast 5:55-72(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=DH484;
 RC MEDLINE=95300974; PubMed=7781779;
 RA Hani J., Stumpf G., Domdey H.;
 RT "PTF1 encodes an essential protein in Saccharomyces cerevisiae, which
 RL shows strong homology with a new putative family of Pplases.";
 FEBS Lett. 365:198-202(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / F11679;
 RA de Haan M., Smits P.H.M., Grievell L.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -I- FUNCTION: ESSENTIAL FOR GROWTH IN SACHAROMYCES CEREVISIAE. MAY BE
 CC INVOLVED IN CYTOKINESIS OR IN CELL SEPARATION.
 CC -I- SIMILARITY: CONTAINS 1 WW DOMAIN
 CC -I- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-21 IS THE INITIATOR.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X85972; CAA59961.1; ALT_INIT.
 DR EMBL; X87611; CAA60941.1; -;

DR EMBL; Z49517; CAA89541.1; -;
 DR PIR; S07867; S07867.
 DR SGD; S0003778; ESSL.
 DR InterPro; IPR000297; -;
 DR InterPro; IPR001202; -;
 DR InterPro; IPR002349; -;
 DR Pfam; PF00639; Rotamase; 1.
 DR Pfam; PF00397; WW; 1.
 DR PRINTS; PR00403; WMDOMAIN.
 DR PROSITE; PS01096; PIC_PPIASE_1; 1.
 DR PROSITE; PS0198; PIC_PPIASE_2; 1.
 DR PROSITE; PS01159; WW_DOMAIN_1; 1.
 DR PROSITE; PS50020; WW_DOMAIN_2; 1.
 KW Isomerase; Rotamase.
 FT DOMAIN 29 62 MW.
 FT DOMAIN 77 190 PIC-LIKE.
 FT CONFLICT 28 28 R -> S (IN REF. 1).
 FT CONFLICT 37 37 V -> A (IN REF. 1).
 FT CONFLICT 147 190 GDLGWFGRGEMQSPFEDAAFLQKYGVSIVSGSGYHVIK
 FT REF -> AASGSGARCSLAKTLIPSSRSAR (IN
 FT REF. 1)
 SQ SEQUENCE 190 AA: 21733 MW: 93E449E3D7BAD989 CXC64;

Query Match 39.4%; Score 367; DB 1; Length 190;
 Best Local Similarity 43.3%; Pred. No. 3.1e-25;
 Matches 77; Conservative 35; Mismatches 50; Indels 16; Gaps 4;

OY 3 STSTGLPPNMTIVSRSHNKEYFLNOSTNESSWDPYGTDEKVLNAYIAFKRNGYKPLV 62
 DB 26 ASRTGLPPNMTIVSRSHNKEYFLNOSTNESSWDPYGTDEKVLNAYIAFKRNGYKPLV 62
 OY 63 NEDGQVRSVSHLLKNNQSRKPKSKSPDGISRTDESIQILKKHLERILSGEYKSLSE--- 119
 DB 75 DHPVAVRCHILIKKDSRRPASHS-ENITISKDAIDELKTLITR-LDDSKNSNEA 132
 OY 120 LANTESSDSSHDRGDLGFESKGMOPPEEAAFNLVGEVSNIIETNSGVHIIORTG 177
 DB 133 LAKERSDCSSYKRGDGLGWFGRGEMQSPFEDAAFLQKYGVSIVSGSGYHVIKRVG 190

RESULT 3
 DOD_DROME 3
 ID DOD_DROME STANDARD: PRT; 166 AA.

AC P54353; O61344; O9VRH1;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE DODO PROTEIN.
 GN DODO OR CG17051.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=OREGON-R; TISSUE=pupae, and Larva;
 RC MEDLINE=96133954; PubMed=8552658;
 RA Maleszka R., Hanes S.D., Hackett R.L., de Couet H.G., Miklos G.L.G.;
 RT "The Drosophila melanogaster dodo (dod) gene, conserved in humans, is
 RL functionally interchangeable with the ESSL cell division gene of
 RT Saccharomyces cerevisiae.";
 Proc. Natl. Acad. Sci. U.S.A. 93:447-451(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=CANTON-S;
 RC MEDLINE=98188272; PubMed=9520435;
 RA Maleszka R., de Couet H.G., Miklos G.L.G.;
 RT "Data transferability from model organisms to human beings: insights
 RL from the functional genomics of the flightless region of Drosophila.";
 Proc. Natl. Acad. Sci. U.S.A. 95:3731-3736(1998).
 RN [3]

RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA MEDLINE-20196006; PubMed-10731132;
 RA Adams M.D., Ceinlker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champs M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abril J.F., Abmayant A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borova D., Botchan M.R., Bouck J., Broxstein P., Brothier P.,
 RA Chertys K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dublin K.D., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laszlo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Splitter E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RT Science 287:2185-2195(2000).
 CC -1- SIMILARITY: BELONGS TO THE PRIC/PARVULIN FAMILY OF ROTAMASES.
 CC -1- SIMILARITY: CONTAINS 1 MW DOMAIN.
 CC CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U35140; AAC46958.1; -;
 DR EMBL: AF017777; AAC28408.1; -;
 DR EMBL: AE003568; AAF50829.1; -;
 DR HSSP: Q13526; 1PIN.
 DR PDBase: FBgn0015379; dod.
 DR InterPro: IPR000297; -;
 DR InterPro: IPR001202; -;
 DR InterPro: IPR002349; -;
 DR Pfam: PF00639; Rotamase; 1.
 DR Pfam: PF00397; MW; 1.
 DR PRINTS: PR00403; WNDOMAIN.
 DR PROSITE: PS01096; PRIC_PPIASE_1; 1.
 DR PROSITE: PS01198; PRIC_PPIASE_2; 1.
 DR PROSITE: PS01159; MW_DOMAIN_1; 1.
 DR PROSITE: PS00202; MW_DOMAIN_2; 1.
 KW Isomerase; Rotamase.
 FT DOMAIN 5 39 MW
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 292 PROTEIN EXPORT PROTEIN PRSA.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT PRIC-LIKE.
 SQ SEQUENCE 166 AA; 18376 MW; 3B4306FA930E7259 CRC64;

Query Match 34.9%; Score 325; DB 1; Length 166;
 Best Local Similarity 41.7%; Pred. No. 1,2e-21;
 Matches 70; Conservative 30; Mismatches 58; Indels 10; Gaps 3;
 QY 8 LPPNWTIRVSRSNKKEFLPNQSTNESSWDPYGTQKEVLNATYAKRNNGYKLVNEDQ 67
 DB 7 LPDGMKRTSRSGMYLMMYKTESQMP--TEP-----AKKRGGSAGGDAPDE 57
 QY 68 VRSHLLINKNOSRKPSKSMKSPDGISTRDESIQIKKHLERITGEVKLSLANTESDC 127
 DB 58 VCHLLVKNHKSRRSRSSWREAN-ITRIKEAQLLEVYKKNKIVQGEAFPFDELARSYDC 116
 DB 117 SSKRGGDLGKFGFRGQMAAFEDAAFLKLVNQLSGIVDSGSHILLR 164
 QY 128 SSHDRGDLGFFSKGQMPPEEAFNLHGEVSNITETNSGVHILQR 175
 DB 117 SSKRGGDLGKFGFRGQMAAFEDAAFLKLVNQLSGIVDSGSHILLR 164
 RESULT 4
 PRSA_BACSU STANDARD; PRT; 292 AA.
 ID P24327;
 DT 01-MAR-1992 (rel. 21, Created)
 DT 01-MAR-1992 (rel. 21, Last sequence update)
 DT 01-OCT-2000 (rel. 40, Last annotation update)
 DE PROTEIN EXPORT PROTEIN PRSA PRECURSOR.
 GN PRSA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1423;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RX MEDLINE-92065824; PubMed-1956302;
 RA Kontinen V.P., Saris P., Sarvas M.;
 RT "A gene (prsa) of *Bacillus subtilis* involved in a novel, late stage
 RT of protein export.";
 RT Mol. Microbiol. 5:1273-1283(1991).
 RL (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-168;
 RA Noback M.A., Terpsstra P., Holsappel S., Venema G., Bron S.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN A LATE STAGE OF PROTEIN EXPORT. PROBABLE
 CC PPIASE.
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE PRIC/PARVULIN FAMILY OF ROTAMASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: X57271; CAA40543.1; -;
 DR EMBL: Y14077; CAA74418.1; -;
 DR EMBL: Z99109; CAB12835.1; -;
 DR PIR: S15269; S15269.
 DR Subtilist; BG10464; prsa.
 DR InterPro: IPR000297; -;
 DR Pfam: PF00639; Rotamase; 1.
 DR PROSITE: PS01096; PRIC_PPIASE_1; 1.
 DR PROSITE: PS01198; PRIC_PPIASE_2; 1.
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 KW Isomerase; Rotamase; Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 292 PROTEIN EXPORT PROTEIN PRSA.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT PRIC-LIKE.
 FT DOMAIN 134 224

SQ SEQUENCE 292 AA; 32510 MW; DEC194DBFE5C9FC0 CRC64;
 Query Match 18.3%; Score 170.5; DB 1; Length 292;
 Best Local Similarity 30.5%; Pred. No. 7.3e-08;
 Matches 51; Conservative 29; Mismatches 46; Indels 41; Gaps 7;
 QY 16 VSRSHKEEYFNNQ-----STNESSWDPPYGTDEKVLNAYIAFKKNGYKPLVNEGQVRV 70
 Db 92 LEKQYKGYLKEQYVETLLTQKADNKKYVDADJKEW-----EGLK-----GKIRA 139
 QY 71 SHLTKNNQSRKPKSWKSPDGISTRDESIQTLKHLERILSGEVKLSLANTESDCSSH 130
 Db 140 SHILVAD-----KKTAEVEVERKKLR-----GE-KFEDLAKYESTDSSA 176
 QY 131 DRGDLGFEFSK-GQMPPFEAFNLHGEVSNILETNSGVHILQRT 176
 Db 177 SKGDLGFMFAKEGQMDFTFSKRAFKLKTGEVSDPVKTOYGYHILIKKT 223
 RESULT 5
 CBE2_CAMJE
 ID CBE2_CAMJE STANDARD; PRT; 273 AA.
 AC Q46105;1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE CELL BINDING FACTOR 2 PRECURSOR (MAJOR ANTIGEN PEB4A).
 GN CBE2 OR PEB4A OR C10596.
 OS Campylobacter jejuni.
 CC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
 CC Campylobacter.
 CC NCBI_TaxID=197;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 55026 / 81-176;
 RX MEDLINE=96099687; PubMed=8525063;
 RA Butuco C., Fremaux C., Pel Z., Tummuru M., Blaser M.J.,
 RA Genatempo Y., Fauchere J.L.;
 RT "Nucleotide sequence and characterization of peba4 encoding an
 RT antigenic protein in Campylobacter jejuni.";
 RL Res. Microbiol. 146:467-476(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCYC 11168;
 RX MEDLINE=20150912; PubMed=10688204;
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
 RA Baaham D., Chillingworth T., Davies R.M., Felwell T., Holtroyd S.,
 RA Jagsels K., Kariyasev A.V., Moule S., Pallen M.J., Penn C.W.,
 RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
 RA Whitehead S., Barrett B.G.;
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
 RT Nature 403:665-668(2000).
 RL Nature 403:665-668(2000).
 CC -1- SIMILARITY: BELONGS TO THE PRIC/PARVULIN FAMILY OF ROTAMASES.
 CC STRONG, TO H.PYLORI HP0175.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X84703; CAAS9175.1; -;
 DR EMBL; AL139075; CAB75232.1; -;
 DR InterPro; IPR000297; -;
 DR Pfam; PF00639; Rotamase; 1;
 DR PROSITE; PS01096; PRIC_PPIASE_1; 1;
 DR PROSITE; PS01098; PRIC_PPIASE_2; 1;
 KW Isomerase; Rotamase; Antigen; Signal.
 FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 273 CELL BINDING FACTOR 2.
 FT DOMAIN 131 228 PPIC-LIKE.
 SQ SEQUENCE 273 AA; 30518 MW; 1C014658BBCD4E39 CRC64;
 Query Match 13.2%; Score 123; DB 1; Length 273;
 Best Local Similarity 28.5%; Pred. No. 0.00092;
 Matches 39; Conservative 24; Mismatches 44; Indels 30; Gaps 6;
 QY 42 DKEVLAATIAFKKNGYKPLVNEGQVRVSHLLKNNQSRKPKSWKSPDGISTRDESIQ 101
 Db 113 DAAKVAFFYDQNKDKYRP-----ARVAKHILVATEERAK-----DITN 152
 QY 102 ILKHLERILSG---EVKLSLANTES-DCSSHDRCGLGFEFSKQMPPEEAFNLHV 157
 Db 153 ELKG-----LKGKELDAFSELAKREKSIDPSKNGGELGWFDDSTWKPFTDAFAALKN 207
 QY 158 GEVSNI-IEFNSGVHIL 173
 Db 208 GTITTPPVKTFNGYHVI 224

RESULT 6
 ID Y175_HELPJ
 ID Y175_HELPJ STANDARD; PRT; 299 AA.
 AC Q9ZMO7;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN JHP0161 PRECURSOR.
 GN JHP0161.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 CC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 CC Helicobacter.
 CC NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonghe B.L., Carmel G.,
 RA Tummino P.J., Carnuso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -1- SIMILARITY: BELONGS TO THE PRIC/PARVULIN FAMILY OF ROTAMASES.
 CC STRONG, TO C.JEJUNI CBE2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE001454; AAD05744.1; -;
 DR HSSP; Q13526; 1PIN.
 DR InterPro; IPR000297; -;
 DR Pfam; PF00639; Rotamase; 1;
 DR PROSITE; PS01096; PRIC_PPIASE_1; 1;
 KW Hypothetical protein; Isomerase; Rotamase; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 299 HYPOTHETICAL PROTEIN JHP0161.
 FT DOMAIN 154 253 PPIC-LIKE.
 SQ SEQUENCE 299 AA; 34040 MW; 9C037BICD1110143 CRC64;

Query Match 13.0%; Score 121.5; DB 1; Length 299;
 Best Local Similarity 29.2%; Pred. No. 0.0014;
 Matches 40; Conservative 23; Mismatches 45; Indels 29; Gaps 6;

```

OY 42 DKEVLNAVIAKFNKNGYKPLVNEGDQVRSVSHLLIKNNOSRKPKSKSPDGSISPTRDESIQ 101
DB 137 EKEMODFYNA-----NKDQLFVKOEAHAR--HILVK-----TEDPAKR 172
OY 102 ILKK-HLERISLGEVKLSLANTES---DCSSHDRCGLGFFSKGOMOPPEEAFLNLAHV 157
DB 173 IISEIDKOPKRAKKEKFEIELARNROTIDPNSKNAONGDLGKFGKNQMAPDFSKAAFAALTP 232
OY 158 GEVSNI-IEITNSGVHIL 173
DB 233 GDYTKTPVKTERTGYHII 249

RESULT 7
Y175_HELPY STANDARD: PRT: 299 AA.
ID Y175_HELPY
AC P56112:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN HP0175 PRECURSOR.
GN HP0175.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
NCBI_TaxID=210;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=26695 / ATCC 700392;
RC MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA Loftus B., Richardson L.M., Lee N., Adams M.D., Hickey E.K.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gokey J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.S., Weidman J.M., Fujii C., Bowman C., Matthey L., Mallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- SIMILARITY: BELONGS TO THE P1C/PARVULIN FAMILY OF ROTAMASES.
CC -1- STRONG, TO C.JEJUNI CBE2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000538; AAD07245.1;
DR TIGR: HP0175;
DR InterPro: IPR000297;
DR Pfam: PF00639; Rotamase; 1.
DR PROSITE: PS01096; P1C_PPIASE_1; 1.
DR PROSITE: PS01098; P1C_PPIASE_2; 1.
KW Hypothetical protein; Isomerase; Rotamase; Signal.
FT SIGNAL 1 21
FT CHAIN 22 299 HYPOTHETICAL PROTEIN HP0175.
FT DOMAIN 154 253 P1C-LIKE.
FT SEQUENCE 299 AA; 34031 MM; E65F3FE94B11F5A CRC64;
SQ
Query Match 13.0%; Score 121.5; DB 1; Length 299;
Best Local Similarity 29.2%; Pred. No. 0.0014;
Matches 40; Conservative 23; Mismatches 45; Indels 29; Gaps 6;
OY 42 DKEVLNAVIAKFNKNGYKPLVNEGDQVRSVSHLLIKNNOSRKPKSKSPDGSISPTRDESIQ 101
DB 137 EKEMODFYNA-----NKDQLFVKOEAHAR--HILVK-----TEDPAKR 172

```

```

OY 102 ILKK-HLERISLGEVKLSLANTES---DCSSHDRCGLGFFSKGOMOPPEEAFLNLAHV 157
DB 173 IISEIDKOPKRAKKEKFEIELARNROTIDPNSKNAONGDLGKFGKNQMAPDFSKAAFAALTP 232
OY 158 GEVSNI-IEITNSGVHIL 173
DB 233 GDYTKTPVKTERTGYHII 249

RESULT 8
P1C_ECOLI STANDARD: PRT: 92 AA.
ID P1C_ECOLI
AC P39159;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE C (EC 5.2.1.8) (PPIASE C)
DE (ROTAMASE C) (PARVULIN).
GN P1C OR PARVA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
NCBI_TaxID=562;
OX [1]
RN SEQUENCE FROM N.A. AND SEQUENCE.
RP MEDLINE=95010704; PubMed=7925971;
RA Rahfeld J.-U., Ruecknagel K.P., Scheldert B., Ludwig B., Hacker J.,
RA Mann K., Fischer G.;
RT "Confirmation of the existence of a third family among
RT peptidyl-prolyl cis/trans isomerases. Amino acid sequence and
RT recombinant production of parvulin."
FEBS Lett. 352:180-184(1994).
RN [2]
RN PRELIMINARY SEQUENCE FROM N.A.
RP STRAIN=K12 / MG1655;
RC MEDLINE=92358234; PubMed=1379743;
RA Daniels D.L., Plunkett G. III, Burland V.D., Blattner F.R.;
RT "Analysis of the Escherichia coli genome: DNA sequence of the region
RT from 84.5 to 86.5 minutes."
Science 257:771-778(1992).
RN [3]
RN REVISIONS, AND IDENTIFICATION.
RP MEDLINE=95184297; PubMed=7878732;
RA Rudd K.E., Sofia H.J., Koonin E.V., Plunkett G. III, Lazar S.,
RA Rouviere P.E.;
RT "Conserved sequence motifs in bacterial and bacteriophage
RT chaperonins."
Trends Biochem. Sci. 20:14-15(1995).
RN [4]
RN SEQUENCE OF 1-21, AND CHARACTERIZATION.
RP MEDLINE=94215709; PubMed=8163020;
RA Rahfeld J.-U., Scherhorn A., Mann K., Fischer G.;
RT "A novel peptidyl-prolyl cis/trans isomerase from Escherichia coli."
FEBS Lett. 343:65-69(1994).
RN [5]
RN FUNCTION: ACCELERATE THE FOLDING OF PROTEINS. IT PREFERS
RN AMINO ACID RESIDUES WITH HYDROPHOBIC SIDE CHAINS LIKE LEUCINE AND
RN CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PEPTIDES SUBSTRATES.
RN -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
RN -1- SIMILARITY: BELONGS TO THE P1C/PARVULIN FAMILY OF ROTAMASES.
RN -----
RN This SWISS-PROT entry is copyright. It is produced through a collaboration
RN between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN the European Bioinformatics Institute. There are no restrictions on its
RN use by non-profit institutions as long as its content is in no way
RN modified and this statement is not removed. Usage by and for commercial
RN entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN or send an email to license@isb-sib.ch).
RN -----
EMBL: S73874; AAB32054.1;
EMBL: M67049; AAB67578.1;
DR

```

DR EMBL; AE000454; AAC76780.1; -.
 DR PIR; S45525; S45525.
 DR EcoGene; EG12352; PPI_C.
 DR InterPro; IPR000297; -.
 DR Pfam; PF00639; Rotamase; 1.
 DR PROSITE; PS01096; PPI_C_PPIASE_1; 1.
 DR PROSITE; PS01098; PPI_C_PPIASE_2; 1.
 DR Isomerase; Rotamase.
 FT INIT_MET 0
 FT SEQUENCE 92 AA; 10101 MW; AE1AE2028277DF4F CRC64;

Query Match 13.0%; Score 121; DB 1; Length 92;
 Best Local Similarity 37.3%; Pred. No. 0.00033;
 Matches 31; Conservative 13; Mismatches 31; Indels 8; Gaps 2;

OY 99 SIOLIKH-----LELLIGSEVLSLANTESDSCSHDGGDGFPSKGMOPPEEA 151
 Db 6 ALHLVKEEKLALDLLEIDKNG-ADFGKLAKKHSICPSGKRGDGEPRGQMWAFDKV 64
 OY 152 AFNLHGEVSNLIETNSGVHILQ 174
 Db 65 VESCVPLEPTGPHGTQFGYHIK 87

RESULT 9
 PINL_HUMAN STANDARD; PRT; 100 AA.
 AC 015428;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PINL-LIKE PROTEIN.
 GN PINL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97446133; PubMed=9299231;
 RA Campbell H.D., Webb G.C., Fountain S., Young I.G.;
 RT "The human PINL peptidyl-prolyl cis/trans isomerase gene maps to human
 RL chromosome 19p13 and the closely related PINL gene to 1p31.";
 RL Genomics 44:157-162(1997).
 CC -1- FUNCTION: NOT KNOWN. COULD BE A TRANSCRIBED PSEUDOGENE.
 CC -1- SIMILARITY: STRONG, TO HUMAN PINL.
 CC -1- SIMILARITY: CONTAINS 1 MW DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U82382; AAB81333.1; -.
 CC HSSP; O13526; 1PIN.
 DR MIM; 602051; -.
 DR InterPro; IPR001202; -.
 DR InterPro; IPR002349; -.
 DR Pfam; PF00397; MW; 1.
 DR PRINTS; PRO0403; MWDOMAIN.
 DR PROSITE; PS01159; MW_DOMAIN_1; 1.
 FT DOMAIN 5 38 MW.
 FT SEQUENCE 100 AA; 11021 MW; SDE04C40C0659A56 CRC64;

Query Match 12.2%; Score 114; DB 1; Length 100;
 Best Local Similarity 36.2%; Pred. No. 0.0015;
 Matches 25; Conservative 8; Mismatches 24; Indels 12; Gaps 1;

OY 8 LPPMWTIRVSRSNKEYELNOSTNESSWDPPYGTDEVLNAYIAKFNKNGYKPLVNEGQ 67
 Db 7 LPPGWEKRSRSPGRGYFNHITNPSQWERPSGN-----SSSGKIMQGEPAR 54
 OY 68 VRVSHLLIK 76
 Db 55 VRSHLLVK 63

RESULT 10
 PPID_HAEN STANDARD; PRT; 622 AA.
 ID PPID_HAEN
 AC P44092;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE D (EC 5.2.1.8) (PPIASE D)
 DE (ROTAMASE D).
 GN PPID OR H11004.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC Haemophilus.
 CC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=RD / KM20 / ATCC 51907;
 RX MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shetty R., Liu L.-T., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus
 RT influenzae Rd.";
 RL Science 269:496-512(1995).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Balroch A.;
 RT Unpublished observations (May-1998).
 CC -1- FUNCTION: PPIASES ACCELERATE THE FOLDING OF PROTEINS. SEEMS TO BE
 CC INVOLVED IN THE FOLDING OF OUTER MEMBRANE PROTEINS (BY
 CC SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
 CC PEPTIDE BONDS IN OLIGOPEPTIDES.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE PPI_C/PARVULIN FAMILY OF ROTAMASES.
 CC -1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; A FRAMESHIFT WAS
 CC INTRODUCED IN POSITION 43 TO PRODUCE THIS ORF.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32781; AAC22655.1; ALT_FRAME.
 DR TIGR; H11004; -.
 DR InterPro; IPR000297; -.
 DR Pfam; PF00639; Rotamase; 1.
 DR PROSITE; PS01096; PPI_C_PPIASE_1; 1.
 DR PROSITE; PS01098; PPI_C_PPIASE_2; 1.
 KW Isomerase; Rotamase; Transmembrane; Cytoplasmic (POTENTIAL).
 FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 17 37 POTENTIAL.
 FT DOMAIN 38 622 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 270 356 PPI_C-LIKE.

SEQUENCE 622 AA: 69590 MW: EE5900A5381C4EB4 CRC64:

Query Match 11.9%; Score 110.5; DB 1; Length 622;
Best Local Similarity 25.7%; Pred. No. 0.033;
Matches 39; Conservative 20; Mismatches 54; Indels 39; Gaps 4;

QY 61 LVNEDGVAVSHLLIKNOSRKPKSKSP-----DGISRT-----R 96
DB 201 LADEMAKOSVDEIKTYEANKSFQPEQYKVOYIDLSDNISRNLQYTVDEIAYQ 260
QY 97 DESIOLKHLERI-----LSGEVLSLANTES-DCSSHDGDLGFFSK 141
DB 261 DNKQFOTQHLAHQFANEODAKVAEELQKGANFADVAKAKSLDKISGENGDGWMNE 320
QY 142 GOMQPFEEAFNLHGEVSNIIETNSGVHIL 173
DB 321 NELPKAFEDAAALQGVQSYQPIWVGNVHIV 352

RESULT 11
NIFM_AZOV1 STANDARD; PRT: 292 AA.
AC P14890:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NIFM PROTEIN.
GN Azotobacter vinelandii.
OS Bacteria; Proteobacteria; gamma subdivision; pseudomonadaceae;
OC Azotobacter.
CC NCBI_TaxID=354;
CC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-69123097; PubMed-2644218;
RA Jacobson M.R., Brigle K.E., Bennett L.T., Setterquist R.A.,
Wilson M.S., Cash V.L., Beynon J., Newton W.E., Dean D.R.;
"Physical and genetic map of the major nif gene cluster from
Azotobacter vinelandii."
RT J. Bacteriol. 171:1017-1027(1989).
CC -1- FUNCTION: REQUIRED FOR THE ACTIVATION AND STABILIZATION OF THE
IRON-COMPONENT (NIFH) OF NITROGENASE. PROBABLE PIIASE.
CC -1- SIMILARITY: BELONGS TO THE PIIIC/PARVULIN FAMILY OF ROTAMASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL: M20568; AAA64732.1;
DR PIR: F32055; F32055.
DR InterPro: IPR000297;
DR Pfam: PF00639; Rotamase; 1.
DR PROSITE: PS01096; PIIIC_PIIASE.1; 1.
DR PROSITE: PS01098; PIIIC_PIIASE.2; 1.
DR Nitrogen fixation; Isomerase; PIIIC-LIKE.
KW DOMAIN 148 243
FT SEQUENCE 292 AA; 32802 MW; 42ED51B1103699BF CRC64;

Query Match 11.3%; Score 105; DB 1; Length 292;
Best Local Similarity 28.2%; Pred. No. 0.037;
Matches 29; Conservative 24; Mismatches 40; Indels 10; Gaps 3;

QY 71 SHLLIKNOSRKPKSKSPDGISRTDESIOILKHLERISGEVLSLANTESDCSSH 130
DB 147 AHLIVTINED-PEENTRE---AARTRIEI-----LKRNGKPEREAFQAKHSECPA 196
QY 131 DRGDLGFFSKGOMQPFEEAFNLHGEVSNIIETNSGVHIL 173

DB 197 MGGILGEEVVPGLYPELDACLFQWARGELSPVLESPIGFHVL 239

RESULT 12
NIFM_AZOV1 STANDARD; PRT: 293 AA.
AC P23119:
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE NIFM PROTEIN.
GN Azotobacter chroococcum mcd 1.
OS Bacteria; Proteobacteria; gamma subdivision; pseudomonadaceae;
OC Azotobacter.
CC NCBI_TaxID=355;
CC [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91358323; PubMed-1885524;
RA Evans D.J., Jones R., Woodley P.R., Wilborn J.R., Robson R.L.;
"Nucleotide sequence and genetic analysis of the Azotobacter
chroococcum nifUSVWZM gene cluster, including a new gene (nifP) which
encodes a serine acetyltransferase."
RT J. Bacteriol. 173:5457-5469(1991).
CC -1- FUNCTION: REQUIRED FOR THE ACTIVATION AND STABILIZATION OF THE
IRON-COMPONENT (NIFH) OF NITROGENASE. PROBABLE PIIASE.
CC -1- SIMILARITY: BELONGS TO THE PIIIC/PARVULIN FAMILY OF ROTAMASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL: M60090; AAA22166.1;
DR PIR: HA3706; HA3706.
DR InterPro: IPR000297;
DR Pfam: PF00639; Rotamase; 1.
DR PROSITE: PS01096; PIIIC_PIIASE.1; 1.
DR PROSITE: PS01098; PIIIC_PIIASE.2; 1.
DR Nitrogen fixation; Isomerase; PIIIC-LIKE.
KW DOMAIN 142 244
FT SEQUENCE 293 AA; 32978 MW; 7D987E2564DE5CDA CRC64;

Query Match 11.1%; Score 103; DB 1; Length 293;
Best Local Similarity 26.7%; Pred. No. 0.056;
Matches 25; Conservative 25; Mismatches 42; Indels 10; Gaps 3;

QY 69 RVSHLLIKNOSRKPKSKSPDGISRTDESIOILKHLERISGEVLSLANTESDCS 128
DB 146 KARHIVTINED-PEENTRE---AARTRIEAI-----LKRNGKPEREAFQAKHSECP 195
QY 129 SHDRGDLGFFSKGOMQPFEEAFNLHGEVSNIIETNSGVHIL 173
DB 196 TMOGILGEEVVPGLYPELDACLFQWARGELSPVLESPIGFHVL 240

RESULT 13
YACD_BACSU STANDARD; PRT: 297 AA.
ID YACD_BACSU
AC P37566:
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 34.1 KDA PROTEIN IN FTSH-CISK INNERGENIC REGION.
GN YACD.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/staphylococcus group; Bacillus.

```

OX NCBL_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
CC -1 SIMILARITY: BELONGS TO THE PPIC/PARVULIN FAMILY OF ROTAMASES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
CC EMBL: D26185; BAA05307.1; -
CC EMBL: Z99104; CAB1848.1; -
CC Subtilist; BG10135; yacd.
CC InterPro: IPR000297; -.
CC Pfam: PF00639; Rotamase; 1.
CC PROSITE: PS01096; PPIC_PPIASE_1; 1.
CC PROSITE: PS50198; PPIC_PPIASE_2; 1.
CC Hypothetical protein; Isomerase; Rotamase.
KW DOMAIN 154 247 PPIC-LIKE.
SQ SEQUENCE 297 AA; 34096 MW; 551B5DEA5E6457B CRC64;

Query Match 10.5%; Score 97.5; DB 1; Length 297;
Best Local Similarity 21.6%; Pred. No. 0.17;
Matches 40; Conservative 36; Mismatches 62; Indels 47; Gaps 7;

QY 14 IYVSRSH-NKEFLNOSTNESSMDPPYGTDE-----VLNAVIAK 52
DB 83 LKISSEVDEFLIKAVANSFYEDHTEKMKDQIRNILLELTDIDISNNELLS 142
QY 53 FNNNGYKPLVNEGDGVRSYHLIKNNQSRKPKSMKSPDISRTDESIOILKKHLERILS 112
DB 143 FYNNK-KELYQFDSDSRHNYVKD-----EEARVLYE---LNG 179
QY 113 GEVKLSELANTESDCSSHDGDLGFFSKQGMOP--FEAAFNLVHGVN-NIETNSG 169
DB 180 GSFSEVAERSTDRYTSYGGDLGFVTEASDNISAYIEAKTLKEDWQSEPIKVSNG 239
QY 170 VHIQ 174
DB 240 VAIQ 244

RESULT 14
BAG3_MOUSE STANDARD; PRT; 577 AA.
AC 09JULV1; Q9JUC7;
DT 01-OCT-2000 (Rel. 40; Created)
DT 01-OCT-2000 (Rel. 40; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE BAG-FAMILY MOLECULAR CHAPERONE REGULATOR-3 (BCL-2 BINDING ATHANOGEME-
3) (BAG-3) (BCL-2-BINDING PROTEIN BIS).
GN BAG3 OR BIS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBL_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20065084; PubMed=10597216;
RA Lee J.H., Takahashi T., Yasuhara N., Inazawa J., Kamada S.,
RA Tsujimoto Y.;
RT "Bis, a Bcl-2-binding protein that synergizes with Bcl-2 in preventing
cell death.";

```

```

RL Oncogene 18:6183-6190(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Osada N., Kusuda J., Tanuma R., Ito A., Hirata M., Sugano S.,
RA Hashimoto K.;
RT "Isolation of full-length cDNA clones from mouse brain cDNA library
made by oligo-capping method.";
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
CC -1 FUNCTION: INHIBITS THE CHAPERONE ACTIVITY OF HSP70/HSC70 BY
PROMOTING SUBSTRATE RELEASE. HAS ANTI-APOPTOTIC ACTIVITY.
CC -1 SUBUNIT: BINDS TO THE ATPASE DOMAIN OF HSP70/HSC CHAPERONES.
CC BINDS TO BCL-2 AND PLC-GAMMA (BY SIMILARITY).
CC -1 PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1 SIMILARITY: CONTAINS 2 WW DOMAINS.
CC -1 SIMILARITY: CONTAINS 2 BAG DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF130471; BAF26840.1; -
CC EMBL: AB041583; BAA95066.1; -
CC InterPro: IPR001202; -.
CC InterPro: IPR003103; -.
CC Pfam: PF02179; BAG; 1.
CC PROSITE: PS01159; WW_DOMAIN_1; 1.
CC PROSITE: PS50020; WW_DOMAIN_2; 1.
CC Chaperone; Apoptosis; Repeat; Phosphorylation.
DR DOMAIN 22 56 WW 1.
DR DOMAIN 126 157 WW 2.
FT DOMAIN 426 503 BAG 1.
FT DOMAIN 528 577 BAG 2.
FT DOMAIN 186 193 POLY-SER.
FT CONFLICT 74 74 N -> D (IN REF. 2).
FT CONFLICT 527 527 P -> Q (IN REF. 2).
FT CONFLICT 539 539 K -> E (IN REF. 2).
SQ SEQUENCE 577 AA; 61828 MW; 367528AC914287E1 CRC64;

Query Match 10.4%; Score 97; DB 1; Length 577;
Best Local Similarity 22.9%; Pred. No. 0.45;
Matches 39; Conservative 29; Mismatches 60; Indels 42; Gaps 8;

QY 2 ASTSTGLPPNWTIRVSRSHNKEFLNOSTNESSMD---PYGTDEKVLNAVIAFKNG 57
DB 18 ASDRDLPPGWEIKIDPQTGMPEFVDHNSRTTWNDRVPPGPKDTASSANGP--SRNG 75
QY 58 YKPLVNEG-----QYRVSHL-----LIKNNQSRKPKSMK--SPDISRTDESIOILKKH 106
DB 76 SRLPLRSGHPYIPOLRPGYIPVPLHSGESNRPPLHFAVSQPGVQRRTAAATTPOR 135
QY 107 LERILSGEVLKSELANTESDCSSHDGDLGFFSKQGM-----OPP 147
DB 136 SOSPLRG--GMTEAAGTQKQ-----GQMPATATTAAGPP 169

RESULT 15
SURA_ECOLI STANDARD; PRT; 428 AA.
ID SURA_ECOLI
AC P21202; P75630;
DT 01-MAY-1991 (Rel. 18; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE SURVIVAL PROTEIN SURA PRECURSOR (PEPTIDYL-PROLYL CIS-TRANS ISOMERASE
SURA) (EC 5.2.1.8) (PPIASE) (ROTAMASE C).
GN SURA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

```

OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RA SEQUENCE FROM N.A.
 RL Almitron M., Tormo A., Kolter R.;
 RP unpublished observations (Jan-1993).
 RN [2]
 RA SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=92334977; PubMed=1630901;
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,
 RA Isono K., Mizouchi K., Nakata A.;
 RT "Systematic sequencing of the Escherichia coli genome: analysis of
 RT the 0-2.4 min region.";
 RL Nucleic Acids Res. 20:3305-3308(1992).
 RN [3]
 RA SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [4]
 RA SEQUENCE OF 331-428 FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=89359108; PubMed=2670894;
 RA Roa B.B., Connolly D.M., Winkler M.E.;
 RT "Overlap between pda and ksa in the complex pda-ksa-apag-apah
 RT operon of Escherichia coli K-12.";
 RL J. Bacteriol. 171:4767-4777(1989).
 RN [5]
 RA SEQUENCE FROM N.A.
 RX Abe S., Okutsu T., Negishi T., Nakajima H., Aono R.;
 RT "n-Hexane sensitivity of Escherichia coli due to low expression of
 RT otaA/omp by insertion of IS2 and identification of the gene product
 RT as an outer membrane protein.";
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RA SEQUENCE OF 21-32.
 RC STRAIN-K12 / EMG2;
 RX MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robinson K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [7]
 RA FUNCTION.
 RX MEDLINE=9030537; PubMed=2165476;
 RA Tormo A., Almitron M., Kolter R.;
 RT "surA, an Escherichia coli gene essential for survival in stationary
 RT phase.";
 RL J. Bacteriol. 172:4339-4347(1990).
 RN [8]
 RA CHARACTERIZATION.
 RX MEDLINE=96198184; PubMed=8626309;
 RA Lazar S.W., Kolter R.;
 RT "SurA assists the folding of Escherichia coli outer membrane
 RT proteins.";
 RL J. Bacteriol. 178:1770-1773(1996).
 RN [9]
 RA CHARACTERIZATION.
 RX MEDLINE=97032152; PubMed=8878048;
 RA Misiakos D., Belton J.W., Raina S.;
 RT "New components of protein folding in extracytoplasmic compartments
 RT of Escherichia coli SurA, FkpA and Skp/OmpH.";
 RL Mol. Microbiol. 21:871-884(1996).
 CC -1- FUNCTION: ASSIST IN THE FOLDING OF EXTRACITOPLASMIC PROTEINS.
 CC ESSENTIAL FOR THE SURVIVAL OF E. COLI IN STATIONARY PHASE.
 CC -1- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
 CC PEPTIDE BONDS IN OLIGOPEPTIDES.

CC -1- SUBCELLULAR LOCATION: PERIPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PPIA/PARVULIN FAMILY OF ROTAMASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: M69521; AAA24304.1; -;
 CC EMBL: D10483; BAA01329.1; -;
 CC EMBL: AE000115; AAC73164.1; -;
 CC EMBL: AB013134; BAA34131.1; -;
 CC PIR: PVO009; PVO009.
 CC Ecogene: EG10985; SurA.
 CC Interpro: IPR000297; -;
 CC Pfam: PF00639; Rotamase; 2.
 CC Prosite: PS01096; PPIA_PPIASE_1; 2.
 CC Prosite: PS01098; PPIA_PPIASE_2; 2.
 CC Kew: Isomerase; Rotamase; Periplasmic; Repeat; Signal.
 FT SIGNAL 1 20
 FT CHAIN 1 428
 FT DOMAIN 171 272 SURVIVAL PROTEIN SURA.
 FT DOMAIN 282 382 PPIA-LIKE.
 FT CONFLICT 25 25 PPIA-LIKE.
 FT CONFLICT 116 116 S -> D (IN REF. 1).
 FT CONFLICT 213 213 G -> GFG (IN REF. 1).
 FT CONFLICT 315 315 K -> E (IN REF. 2).
 FT CONFLICT 325 325 A -> T (IN REF. 2).
 FT CONFLICT 333 333 G -> V (IN REF. 2).
 SQ SEQUENCE 428 AA; 47283 MW; 25F6ADAB903CDB8E CRC64;

Query Match 9.58; Score 89; DB 1; Length 428;
 Best Local Similarity 26.4%; Pred. No. 1.5;
 Matches 29; Conservative 23; Mismatches 42; Indels 16; Gaps 4;
 QY 67 QVRSHTLLNNQSRKRSKSPGDSRTDESIQI-LKKHLERILSGEYKLSPLATES 125
 DB 284 EVNARHILK-----BSPT--MTDEQARVLEQIADIRKSGKTTFAAAAEFS 329
 QY 126 -DCSSHDGRGDLGFFSKQMQPPEEAFLHVGESVNIETNSGVHILQ 174
 DB 330 QDPESANQGGDLGATPDIDPAPRALTRLNKGQMSAPVHSSFGWHLIE 379

Search completed: June 8, 2001, 09:39:14
 Job time: 79 sec

This Page Blank (uspio)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2001, 04:01:10 ; Search time 1501.23 Seconds
(without alignments)
9715.788 Million cell updates/sec

Title: US-09-507-242-1
Perfect score: 989
Sequence: 1 gatcaaccacatagatgtgtg.....ggtttgtatattgtgattcc 989

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1283235 seqs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: gb_pl5:*
17: gb_ba2:*
18: gb_fun:*
19: gb_hggo_hum:*
20: gb_hggo_inv:*
21: gb_hggo_rod:*
22: gb_hggo_hum1:*
23: gb_hggo_hum2:*
24: gb_hggo_hum3:*
25: gb_hggo_hum4:*
26: gb_hggo_hum5:*
27: gb_hggo_hum6:*
28: gb_hggo_hum7:*
29: gb_hggo_hum8:*
30: gb_hggo_hum9:*
31: gb_hggo_hum10:*
32: gb_hggo_hum11:*
33: gb_hggo_hum12:*
34: gb_hggo_hum13:*
35: gb_hggo_hum14:*
36: gb_hggo_hum15:*
37: gb_hggo_hum16:*
38: gb_hggo_hum17:*
39: gb_hggo_hum18:*
40: gb_hggo_hum19:*
41: gb_hggo_hum20:*
42: gb_hggo_hum21:*
43: gb_hggo_hum22:*

44: gb_ov:*
45: gb_pat:*
46: gb_ph:*
47: gb_pl:*
48: gb_in:*
49: gb_om:*
50: gb_ov:*
51: gb_pat:*
52: gb_ph:*
53: gb_pl:*
54: gb_in:*
55: gb_om:*
56: gb_ov:*
57: gb_pat:*
58: gb_ph:*
59: gb_pl:*
60: gb_in:*
61: gb_om:*
62: gb_ov:*
63: gb_pat:*
64: gb_ph:*
65: gb_pl:*
66: gb_in:*
67: gb_om:*
68: gb_ov:*
69: gb_pat:*
70: gb_ph:*
71: gb_pl:*
72: gb_in:*
73: gb_om:*
74: gb_ov:*
75: gb_pat:*
76: gb_ph:*
77: gb_pl:*
78: gb_in:*
79: gb_om:*
80: gb_ov:*
81: gb_pat:*
82: gb_ph:*
83: gb_pl:*
84: gb_in:*
85: gb_om:*
86: gb_ov:*
87: gb_pat:*
88: gb_ph:*
89: gb_pl:*
90: gb_in:*
91: gb_om:*
92: gb_ov:*
93: gb_pat:*
94: gb_ph:*
95: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	534	54.0	534	13	AF224269
2	99.8	10.1	865	53	AF429018 clone BAO
3	97	9.8	895	12	AF035768 Emericell
4	94.2	9.5	836	6	DD078757
5	81.8	8.3	987	53	CNS06N09
6	76.8	7.8	951	53	CNS06DSQ
7	72.2	7.3	707	14	DLA133755
8	65	6.6	7218	10	166494
9	64.8	6.6	549	14	NCR6023
10	61.2	6.2	954	94	AB009691
11	58	5.9	87947	12	AC006201

C 12	57.4	5.8	529	8	AF239760	Xenopus 1
C 13	56.6	5.7	1023	14	SCYR017C	249517 S.cerevisia
C 14	56.6	5.7	2689	14	SCPEF1	X85972 S.cerevisia
C 15	56.6	5.7	43661	14	SCXCOSM83	X87611 S.cerevisia
C 16	54	5.5	993	6	DMU35140	U35140 Drosophila
C 17	53.6	5.4	994	93	HSU49070	U49070 Human pepti
C 18	53.6	5.4	1014	9	AR074009	AR074009 Sequence
C 19	53.6	5.4	1014	9	AR082120	AR082120 Sequence
C 20	52.2	5.3	936	54	CNS07ANM	AL436808 t7 end of
C 21	51	5.2	44751	15	SPCC16C4	AL031535 S.pombe c
C 22	48.8	4.9	38193	65	AC017671	AC017671 Drosophila
C 23	48.8	4.9	298640	2	AE003467	AE003467 Drosophila
C 24	48.6	4.9	981	2	BS33KDA	X57271 B.subtilis
C 25	48.6	4.9	10617	2	BSY14077	Y14077 Bacillus su
C 26	48.6	4.9	210440	2	BSUB0006	Z91009 Bacillus su
C 27	48	4.9	220808	65	AC018700	AC018700 Homo sapi
C 28	45.8	4.6	186897	73	AC067953	AC067953 Homo sapi
C 29	45.2	4.6	98734	84	PFMALP2	AL031745 Plasmodin
C 30	45	4.6	150561	64	AC016457	AC016457 Homo sapi
C 31	45	4.6	299550	2	AP001511	AP001511 Bacillus
C 32	44.8	4.5	59320	4	AC024746	AC024746 Caenorhab
C 33	44.8	4.5	155711	6	YIYD10	AL360354 Plasmodin
C 34	44.8	4.5	180664	60	AC006706	AC006706 Caenorhab
C 35	44.4	4.5	405	2	AF247686	AF247686 Francisel
C 36	44.4	4.5	195969	76	AC074354	AC074354 Oryza sat
C 37	44.2	4.5	4349	5	AF286897	AF286897 Plasmodin
C 38	44.2	4.5	165213	76	AC074367	AC074367 Homo sapi
C 39	43.8	4.4	16422	4	AE001421	AE001421 Plasmodin
C 40	43.6	4.4	147901	80	AL355587	AL355587 Homo sapi
C 41	43.2	4.4	916	63	CNS06090	AL407796 t3 end of
C 42	43.2	4.4	169794	60	AC004688	AC004688 Plasmodin
C 43	43	4.3	42043	5	CELC33E10	AF067220 Caenorhab
C 44	43	4.3	104992	60	AC005504	AC005504 Plasmodin
C 45	43	4.3	169546	60	AC004157	AC004157 Plasmodin

ALIGNMENTS

RESULT 1	AF224269	534 bp	DNA	PLN	02-FEB-2001
LOCUS	AF224269				
DEFINITION	Candida albicans peptidyl prolyl cis/trans isomerase gene, complete cds.				
ACCESSION	AF224269				
VERSION	AF224269.1	GI:12655855			
KEYWORDS	Candida albicans.				
SOURCE	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales; anamorphic Saccharomycetales; Candida.				
REFERENCE	1 (bases 1 to 534)				
AUTHORS	Devasahayam,G., Chaturvedi,V. and Hanes,S.D.				
TITLE	Isolation of the Candida albicans homolog of the ESS1 gene, encoding a prolyl isomerase				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 534)				
AUTHORS	Devasahayam,G., Chaturvedi,V. and Hanes,S.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-JAN-2000) Biomedical Science, Wadsworth				
FEATURES	Center/SUNY-Albany, 120 New Scotland Avenue, Albany, NY 12208, USA				
SOURCE	Location/Qualifiers				
1.	534				
/organism="Candida albicans"					
/db_xref="taxon:5476"					
<1..>534					
/product="peptidyl prolyl cis/trans isomerase"					
1.	534				
/note="ESS1: Isolated by complementation of Saccharomyces cerevisiae ess1-ts mutant"					
/codon_start=1					
/trans_table=12					
/product="peptidyl prolyl cis/trans isomerase"					

/protein_id="AAK00626.1"
/db_xref="GI:12655856"
/translation="MASTSTGLPNNWIRVSRSHNKEFLINOSTNESSMDPPYGTDRKE
VLNAVYIAKFNKNGCYKPLVNEDDGQVRSYHLIKNNOSRRKPSKSPGIISTRDESIOI
LKKHLRIISGKELSELIANTESDCSHDGGDLGFFSKGQMGPPPEAAFNHVEV
SNIEFNCSVHIIQRTG"

BASE COUNT 192 a 95 c 114 g 133 t
Query Match 54.0%: Score 534; DB 13; Length 534;
Best Local Similarity 100.0%: Pred. No. 3.3e-99;
Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	256	atggcatcgacatacaagaagcttaccacctaattgacgatatgaagatccagatcccat	315		
Db	1	ATGGCATCGACATCAACAGCGTTACCACTATTTGAGACGATTACAGTATCCAGTCCAT	60		
Qy	316	aacaagaagatattctttaaaccaatctacaaatgagtcgctcttgagcccaactatggc	375		
Db	61	AACAAAGAGATTTCTTAAACCAATCTACCAATGATGCTCTTGGAACCCACTTATGGC	120		
Qy	376	actgacaaagaagfatggaatgcatatctgcgaagtttaaaacaatggttaacagcca	435		
Db	121	ACTGACAAAGAGATTTGATTCATTCATTCGCAAGTTTAAAAACAATGGTTACACCCA	180		
Qy	436	cttgatgataggatgagccaggttagagttctcatcttgatgatacaacaatcaatca	495		
Db	181	CTTGATGATGAGATGAGCGCAGGTAGAGTTCTCATTTGTTGATCAAGAACAAATCA	240		
Qy	496	agaaaacccaagcttctggaagtcctcccgatggttaagtgaactaagaagcgaatcata	555		
Db	241	AGAAAACCCAAGCTTCTGGAAGTCCCGAGTGTAAAGTAACTAGACAGCAATCTATA	300		
Qy	556	cagatatggaagaacatttggaaagaatatgagttggtgagttaaactaagtgaatg	615		
Db	301	CAGATATTGAAGAAACATTTGGAAAGATATTTGAGTGTGATTAACATAAGCATTTG	360		
Qy	616	gcaaatccgaagaatgattgagctcacatgacagagtggtgatttaagggttttttaagc	675		
Db	361	GCAAAATCCGAAGATGATTGCAGCTCACATGACAGAGTGTGATTAAGGCTTTTAAAGC	420		
Qy	676	aaaggaacaatgacacacatcttgaagaagccgcatcatttaatttgatggttggaagtc	735		
Db	421	AAAGGACAAATGACCAACCACTTGAAGAACCCGCAATTCATTTGCAATGGTGAAGATC	480		
Qy	736	agtaacaatggaacacatagtggtgtccatatccctccaagaagaagataa	789		
Db	481	AGTAACATTAATTGAACCAATAGTGCTCATATCTCTCCAAAGACAGATTA	534		
RESULT 2	CNS074N8	865 bp	DNA	SNS	11-JAN-2001
LOCUS	CNS074N8				
DEFINITION	clone BA0AB029808 of library BA0AB from strain CLIB 210 of Kluyveromyces lactis, sequence tagged site.				
ACCESSION	AL429018				
VERSION	AL429018.1	GI:12212212			
KEYWORDS	STS.				
SOURCE	Kluyveromyces lactis.				
ORGANISM	Kluyveromyces lactis.				
REFERENCE	1 (bases 1 to 865)				
AUTHORS	Boloiteau-Fukuhara,M., Toffano-Nicche,C., Attiguenave,F., Duchateau-Nguyen,G., Lemaire,M., Marneisse,R., Montocher,R., Robert,C., Termler,M., Wincker,P. and Wesolowski-Louvel,M.				
TITLE	Genomic Exploration of the Hemiascomycetous Yeasts: 11.				
JOURNAL	Kluyveromyces lactis				
PUBMED	FEBS Lett. 487 (1), 66-70 (2000)				
REFERENCE	2 (bases 1 to 865)				
AUTHORS	Souciet,J.L., Aigle,M., Attiguenave,F., Blandin,G.,				

TITLE
JOURNAL PUBMED
REFERENCE
AUTHORS
JOURNAL
COMMENT
FEATURES
SOURCE
misc_feature
misc_feature
misc_feature
misc_feature
BASE COUNT
ORIGIN
Query Match
Best Local Similarity
Matches 202; Conservative 0; Mismatches 127; Indels 6; Gaps 2;
456 ggttagagtttcattggtgatcgaagaacaatcaatcaagaagaacccaagttctgaa 515
165 GGTGAGACCTTGGATCTCTGATTAAACACAGAGATTCTGAGAGACCAAGCTC---CA 221
516 gtccccaagatgtaagtagaactagagagaactctatagatatatgaagaacaactt 575
222 TAGAAATGAAATATTTACATTACTTACGATGAGAGGCTTAAGACAGATTAAGAACTTATAT 281
576 ggaagaataatctagtggtgaggtctaaactaagtaactggaatctggaacaaaccgaagtgtg 635
282 CAAAGAG---ATTGAATGTTGGTGAACCACTTGGATTCATTGGCTTAAGAGAGAGATGATTG 338
636 caagctacatgacagaggtggtgatttagggtttttagcaagaagacaatgcaaccacc 695
339 TTTCATCAGCGAAAGAGAGTGTGATTTGGGGTCTTTCGGTCATGAGAAATGCAACCTTC 398
696 attcgaagaagccgcatcctaattgcactgttggaggaagtcagtaatacataattgaaccac 755
399 TTTCGAAAGAGCTTCGTTGGCTTGAATAATGCATCAAGTCAAGTATATTTGAGAAATCTGA 458

deMontigny, J., Dujon, B., Durieux, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuvéglise, C., Ozler, K., Kalogeropoulos, O., Potier, S., Saurin, W., Tekala, F., Toftano, N., Nioche, C., Weslowski, L., Douvel, M., Winkler, P., and Weissenbach, J.
Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of yeast species for molecular evolution studies(1)
FEMS Lett. 487 (1), 3-12 (2000)
1152876
3 (bases 1 to 865)
Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seque@genoscope.cns.fr Web : www.genoscope.cns.fr)
This STS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvatum*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
Location/Qualifiers
1. 865
/organism="Kluyveromyces lactis"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone="BA0AB0295B0"
/clone_id="BA0AB"
1. >489
/note="similar to *Saccharomyces cerevisiae* ORF YJR017C [ES1 : processing/termination factor 1]"
/evidence=not experimental
complement(<12..>236)
/note="similar to *Saccharomyces cerevisiae* ORF YJR018w [questionable ORF]"
/evidence=not experimental
complement(<506..>865)
/note="similar to *Saccharomyces cerevisiae* ORF YGR203w [weak similarity to X.laevis protein-tyrosin-phosphatase cdc homolog 2 and to hypothetical protein YPR200c]"
/evidence=not experimental
265 a 165 c 192 g 243 t

Oy	756	taatggtgcctatatccccaagaagacggataaa	790
Dd	459	cagtgccttcattcatcattcaaaaagacttccctga	493
<hr/>			
RESULT	3		
LOCUS	AF035768	895 bp	mRNA PLN 27-MAR-1998
DEFINITION	Emeritella nidulans peptidyl-prolyl cis/trans isomerase (plna)		
ACCESSION	AF035768		
VERSION	AF035768.1 GI:2739196		
SOURCE	Aspergillus nidulans.		
ORGANISM	Aspergillus nidulans Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emeritella. 1 (bases 1 to 895) Crenshaw,D.G., Yang,J., Means,A.R. and Kornbluth,S. The mitotic peptidyl-prolyl isomerase, plnI, interacts with Cdc25 and Plx1 EMBO J. 17 (5), 1315-1327 (1998)		
JOURNAL	MEDLINE 98151356		
REFERENCE	2 (bases 1 to 895) Crenshaw,D.G., and Means,A.R. AUTHORS Direct Submission TITLE Submitted (25-NOV-1997) Pharmacology and Cancer Biology, Duke University Medical Center, LSRC Bldg, Lasalle Street Extension, Durham, NC 27708, USA		
FEATURES	Location/Qualifiers 1..895 /organism="Aspergillus nidulans" /strain="R153" /db_xref="taxon:5072" 1..895 /gene="plna" 55..585 /gene="plna" /note="PINA: parvulin type" /codon_start=1 /product="peptidyl-prolyl cis/trans isomerase" /protein_id="AAC49984.1" /db_xref="GI:2739197" /translation="MVTGGLPAGWEVRHSKSNLPYFNPATRESNRPADDMETL KMVAITYHSAATYHEAPSQEGKIRCSHLVTKRDSRPPSWREAEITRKKEAREIL RGHERIMGEIRLDLAISESDCSSARKKGDLQFGRGEWQKEFEAAALAPCGOVS DIVSGSGHLIERLO"		
BASE COUNT	252 a	208 c	217 g 218 t
ORIGIN			
Query Match	9.8%	Score 97:	DB 12: Length 895;
Best Local Similarity	51.8%;	Pred. No. 3.7e-10;	
Matches 273;	Conservative 0;	Mismatches 245;	Indels 9; Gaps 2;
Oy	271	acaaggctaacacctaatgtgacgatitagatgatlaccagatcccataaagaatlatlc	330
Dd	64	ACCGGTCTCCAGCAGGAGGTGGGAAGTTGCCATTGCAGACTCCAAAAAACCCTTCCTACTAC	123
Oy	331	ttaaaccaatctaccaatgatgctgcttggaagccaccttatgtgcactgcagaagaagta	390
Dd	124	TTCACAACCCCGTTACCAGAGCATCACGATGGCAACCTCCAGCGGATACCGATATGAGACT	183
Oy	391	ttagatgcatactatcggaagttaaaaaaatg-----ttacaagccacttgtgat	444
Dd	184	CTTAGATGTACATGGCACATATCACAGCGCGCCGCTACTTACCATGAAGCCCCGAGT	243
Oy	445	gaggatgcccaggtctaagtttcacattgtgtgaataagaacatacaataagaanaacc	504
Dd	244	CAGAAGAGGCAAGATCCGTTGCAGTCACCTCTCTACTCAAAGCAGAGAGAGGCGACCG	303
Oy	505	aagctctggaagctccccaagatgatatagaatgaactgaagaactgaactatacagatattg	564

```

Db      304 AGCAGCTGGAGGAGCGGAG--ATTACTGGAACAAGAGGAGCGCGGAGATTGTTG 360
Qy      565 aagaacattggagaagaataatgagtgtgagttaaactaaagttatggaataacc 624
Db      361 CGCGGTCAATGAGGAGCAATATCGAGGAGATCCGGCTCGGAGATCTTGCCATGTC 420
Qy      625 gaaagtattgacgtacacatgacagaggtgtgattgaaggttttttaagcaagaca 684
Db      421 GAGTCGAGACTGCACTAGTGTAGAGAGAAGGCGATCTCGGCTTCTTTGGCGGTGAA 480
Qy      685 atgaaccacacattcgaagaagccgcatcattgtcatgttgaagaatgaacata 744
Db      481 ATGCGAAGAAAGCTTCGAGGAGGAGCGCTTCGCAATTGCAACCGGGTCAAGTACGACATT 540
Qy      745 atgaaccacatgagtgtgtccatcatccccaagaacaggaataat 791
Db      541 GTCGATTCAGGGTCTGCTGCTTCATCTCATCGAGCGGCTACATAGAT 587

RESULT 4
LOCUS   DD078757      836 bp      mRNA      INV      27-NOV-1996
DEFINITION Dictyostelium discoideum Pina (pina) mRNA, partial cds.
ACCESSION U78757
VERSION   U78757.1 GI:1688321
KEYWORDS   .
SOURCE      Dictyostelium discoideum.
ORGANISM   Dictyostelium discoideum.
REFERENCE   Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
AUTHORS     Loomis,W.F. and Iranfar,N.
TITLE       Direct Submission
JOURNAL     Submitted (19-NOV-1996) Department of Biology, University of
            California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0022,
            USA

FEATURES
    source          location/Qualifiers
    gene            1..836
                        /organism="Dictyostelium discoideum"
                        /strain="AX4"
                        /db_xref="taxon:44689"
                        <1..733
                        /gene="pina"
                        /gene="pina"
                        <1..733
                        /note="similar to human Pina1 encoded by the sequence
                        presented in GenBank Accession Number U49070"
                        /codon_start=2
                        /product="Pina"
                        /protein_id="AAB36960.1"
                        /db_xref="GI:1688322"
                        /translation="PEWASKRPIENVLEIKNGVINDRVIDSKDFTVGRSSEVASV
                        LIDHPVSRRHALVYGANNRFLYLDIOAGTGOVNDQVKPLPTTVENETPFKG
                        SSKHFKITGVYNNPSSSSSSSSSSSEKTYTCRLILYKHGSRPSSWRBSKTR
                        TKERAIKLMEYRATIISSGATREDLHAKNSDCSARKGITLDPFKGQMRPEDCA
                        FSKVEVSSIVPDSGVHIERLA"
    CDS            121 c 127 g 270 t

BASE COUNT      318 a 121 c 127 g 270 t
ORIGIN
Query Match          9.58; Score 94.2; DB 6; Length 836;
Best Local Similarity 57.48; Pred. No. 1.4e-09;
Matches 190; Conservative 0; Mismatches 138; Indels 3; Gaps 1;

```

```

Db      532 ATCTGCTTCGCCCATTTGAAAGATTAGCTCATTAATAATTCAGATGTGACAGCTGCTAA 591
Qy      648 cagagtggtgattagaggttttttagcaagagacatgaaccacatcgaagac 707
Db      592 AAGAGTGGCTATCTTATTCATTCATCAAAAGAGGCCAATGCAAGACCTTTTGAAGATTG 651
Qy      708 cgcattcaattgcatgttggagaagtcagtaacataattgaaccaatagtggtcca 767
Db      652 CGCATTTCTTTAAAGTTGGTGGATGAGTATGATGATACGATTCAGAGGTGTTCA 711
Qy      768 tatctccaagaacagataatcaagata 798
Db      712 TATTATTGAAGATTAGCATTAATAATGTAATA 742

RESULT 5
LOCUS   CNS06N09      987 bp      DNA      STS      10-JAN-2001
DEFINITION r3 end of clone AU0AA008E01 of library AU0AA from strain CBS 3082
ACCESSION AL406159
VERSION   AL406159.1 GI:12169821
KEYWORDS   STS.
SOURCE      Saccharomyces kluyveri.
ORGANISM   Saccharomyces kluyveri.
REFERENCE   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetes kluyveri.
AUTHORS     Neuveglise,C., Bon,E., Lepingle,A., Wincker,P., Artiguenave,F.,
            Galliardin,C. and Casaregola,S.
TITLE       Genomic Exploration of the Hemiascomycetous Yeasts: 9.
JOURNAL     Saccharomyces kluyveri
PUBMED     FEBS Lett. 487 (1), 56-60 (2000)
11152884
2 (bases 1 to 987)
Souchet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boletoin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,
de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
Malpertuy,A., Neuveglise,C., Ozler-Kalogeropoulos,O., Potter,S.,
Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
Wincker,P. and Weissenbach,J.
Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
yeast species for molecular evolution studies(1)
FEBS Lett. 487 (1), 3-12 (2000)
11152876
3 (bases 1 to 987)
Genoscope.
Direct Submission
Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail:
segref@genoscope.cns.fr - Web: www.genoscope.cns.fr)
This STS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
location/Qualifiers
    source          1..987
                        /organism="Saccharomyces kluyveri"
                        /strain="CBS 3082"
                        /db_xref="taxon:4934"
                        /clone="AU0AA008E01"
                        /clone_lib="AU0AA"
                        /note="end : T3"
                        <365..>838
                        /note="similar to Saccharomyces cerevisiae ORF YB017c.1
                        ESI1 ; processing/termination factor 1 ]"
    misc_feature

```



```

/codon_start=1
/product="peptidyl-prolyl cis-trans isomerase"
/protein_id="CAB94994.1"
/db_xref="GI:8670993"
/translation="MSSEKVRASHIILKHGSRKRSKMDPDGSLISATITDDAVSOL
QSLQELISDPASFSDLASRSHGSSKAKRGGLAPFGQGMQKFFEEATFALAKVEIS
DIVYDGEVHLIKRTG"
BASE COUNT      188 a      185 c      162 g      172 t
ORIGIN

```

Query Match	7.38;	Score 72.2;	DB 14;	Length 707;
Best Local Similarity	66.2%;	Pred. No. 4.2e-05;		
Matches 104;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;

O7 632 atgycagccacatgacagagtgttatttagtglttttagcaaaagacaatatgcaac 691
 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 313 ACTGCAGCTCTGCAAAAGCGTGTGGCATCTTGCCPTTTGGAAGGGGCCAGATCAGA 372

Oy 652 caccattcgaagaagccgcattcaatttcgacgttgggaaagtcagttaacataatctgaaa 751
 ||| | | | | | | | | | | | | | | | | | | | | |
Db 373 AACCTTTCGAAGAACCCACATTTGCACATAAAGCTTGCTGAGATAAGTGATATCGTGAGATA 432

```
QY      752   coaatagtgtygtccatatcctccaagaacagata    788
          | | | | | | | | | | | | | | | | | |
Db     433   CTGACAGTGGAGTTCAATCATCAAGAGAACAGGATA    469
```

RESULT		8	
I66494/C			PAT
LOCUS	I66494	7218 bp	DNA
DEFINITION	Sequence 14 from patent US 5670367.		
ACCESSION	I66494		
VERSION	I66494.1	GI:2724471	
KEYWORDS	.		
			28-DEC-1997

ORGANISM	Unknown.
REFERENCE	Unclassified.
AUTHORS	1 (bases 1 to 7218)
TITLE	Doerner, F., Schefflinger, F. and Falkner, F. Gunter
JOURNAL	Recombinant fowlpox virus
FEATURES	Patent: US 5670367-A 14 23-SNP-1997;
SOURCE	Location/Qualifiers
	1..7218

BASE COUNT	1944 a	1491 c	1486 g	1929 t	368 others
ORIGIN					

Query Match	6.68; Score 65; DB 10; Length 7218;
Best Local Similarity	2.1%; Pred. No. 0.001;

Matches 8; Conservative 232; Mismatches 137; Indels 0; Gaps 0;

OY 525 tggatataagtagaacctagagacgaattctatacagatatgtgaagaaacaattgtaaaaagaat 584
 ||||| :::: :::: : : : : : : : : : : : : : : : :
Db 1440 TCGTACRRR 1380

[illegible]

OY 645 tgcacagaagtggtgtaattgggtttttaacaaagacaatgcaaccaccttgaaga 704
 :: : ::
Db 1320 rr 1261

Oy 705 agccgcattcaatttgcattgttggaagtcagttaacataattgaaaccaatgltgtgt 764
::
Db 1260 RR 1300

```

Oy 765 ccatactcacaagacagataaacaagatatgtgattgataaanaatgaaata 824
      : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1200 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1141

```

Qy	885	actgtgagagaacatc	901
Db	1080	RRRRRRRRRRRRRATC	1064

RESULTS	9	NCRC023	E40	SPRM	DTN	20-NOV-1000
---------	---	---------	-----	------	-----	-------------

LOCUS	NC00023	345 bp	1873
DEFINITION	Neurospora crassa mRNA for parvulin (sspl).		
ACCESSION	AJ006023		
VERSION	AJ006023.1	GI:3127914	
KEYWORDS	parvulin; peptidylprolyl isomerase; sspl gene.		
SOURCE	Neurospora crassa.		
ORGANISM	Neurospora crassa		

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 549)
Tropschug, M.
Direct Submission
Submitted (06-MAY-1998) Tropschug M., Institut für Biochemie und
Molekularbiologie, Universität von Freiburg, Hermann-Herder-Str 7,
78122 Freiburg, Germany

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE

2 (bases 1 to 549)
Kops,O., Eekelskorn,C., Hottenrott,S., Fischer,G., Mi,H. and
Tropschug M.
Ssp1, a site-specific parvulin homolog from *Neurospora crassa*
active in protein folding
J. Biol. Chem. 273 (48), 31971-31976 (1998)
99041963

FEATURES	
source	Location/Qualifiers
	1..549
	/organism="Neurospora crassa"
	/db_xref="taxon:5141"
	/cell_line="74 A"
gene	1..549
	/gene="sp1"
CDS	1..549
	/gene="sp1"

```

/gene= ssp1
/EC_number="5.2.1.8"
/note="parvulin"
/codon_start=1
/evidence=experimental
/product="peptidylprolyl isomerase"

```

```

/protein_id="CA06818.1"
/db_xref="GI:3127915"
/db_xref="SPTREMBL:O60045"
/translation="MSNTIERGLPEDMVRHSOSKNLPYENSATKTSRMPEPGTDV
DKLTIYAKVYHSPSCOOOOOOPGKICALLVHNSRPPSSRSEELITRTKK
EALTTDGLGFEGRISGCSISIELALTEBDCSSAKKRGLDGLFGKDMKREFEDAAFA
KPEISDGLVDVPAISLHLERLE"
BASE COUNT      137 a      174 c      145 g      93 t
ORIGIN

```

Query Match	6.68;	Score 64.8;	DB 14,	Length 549;
Best Local Similarity	48.58;	Pred. No. 0.0014;		
Matches 254; Conservative	0;	Mismatches 252;	Indels 18;	Gaps 2;

QY 271 acagcctaccacctaattgacgatagatcatccagatccatacaaaagtatttc 350
|| || | ||| | || |||| | ||
Db 19 ACCGGTCTCCCCGAGACACTGGGAGATCCGCCACTTCGAATCCAGAAGACCTGCCCTACTAC 78

QY 331 ttaaccaatctaccaatgagtcgtcttggaccacattatgcaactgacaagaagta 390
|| ||| | ||||| || ||||| ||||| ||||| ||||| ||
Db 79 TTCAACAGCGCCACCACAAGACTCTTCGCTGGGAGCACCACTTCGGGACACCGAGCCTTGACAA 138

```

QY      391 ttgatgcatacatctgcygaatt-----taaaacatggtttacaagcca 435
      | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      139 CTCAGATCTTACATATGCGCCAAATGACCACTTCGCCACGATCAACAGCAGCAACAACACAG 198

```

Oy	436	ctctgtgaatctgaggtctgcgaagttaaggtcttcacattctgtatctgaacaagaacatca	495
Db	139	CAGCAACACACCAAGGGAAATAGCTGTCCGCACTTTGGCTCAAGCAACCAAGC	238
Oy	496	agaaaccacaagctctctggaatctcccaagatgctctaaagtgaactaagaacgaatcata	555
Db	259	AGGAGGCCCACTTTTGTGGCGGGAATCCGA---GATTACCCGACACCAAGCAAGAGCCCTC	315
Oy	556	cagaatattgaagaacaactcttgaagaataattggaatggttgagtctaaactaagtgaattg	615
Db	316	ACTACCCGCAAGGATTCGAAACACGCGATCAAGTCGGCGTCACATTCCTCGGGAGACTG	375
Oy	616	gcaaataccgaagtgattctgacgtcacatgacagaggtggtgatttaaggtcttttagc	675
Db	376	GCCCTGACCGAGTCCGCACTGCTCTCTGCTCGAAGCGGGGTGACCTCGGCTACTTGGGA	435
Oy	676	aaagaacaaatgcaaacaccactcttgaagaagccgcatlcaattgtcgtcttgagaagtc	735
Db	436	CGCGCGGCACATGCAAAAGAGACTTTGAGAGATGCCCTTTTCCGTTGAAGCCCTGGCGAGATC	495
Oy	736	agtaacaaatctgaacacaatagtgtgtctcatatctctccaag	79
Db	496	AGCGACATCTGGACACGCGCTAGTGGGTTGCATTTGATTGAACG	539

```

RESULT 10
AB009691 LOCUS 26-NOV-1999
DEFINITION Mus sp. mRNA for PIN1, complete cds.
ACCESSION AB009691
VERSION AB009691.1 GI:6468199
KEYWORDS mpin1; PIN1.
SOURCE Mus sp. cDNA to mRNA.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE Fujimori,F., Takahashi,K., Uchida,C. and Uchida,T.
TITLE Mice lacking Pin1 develop normally, but are defective in entering
cell cycle from G(0) arrest
BIOCHEM. BIOPHYS. RES. COMMUN. 265 (3), 658-663 (1999)
JOURNAL 20070807
MEDLINE 2 (bases 1 to 954)
REFERENCE Fujimori,F. and Uchida,T.
AUTHORS Direct Submission
TITLE Submit (12-DEC-1997) to the DDBJ/EMBL/Genbank databases.
JOURNAL Fumihito Fujimori, Science University of Tokyo, Department of
Biological Science & Technology, Laboratory of Genome Biology;
Yamazaki, Noda, Chiba 278, Japan
(E-mail:fujimorf@res.noda.sut.ac.jp, Tel: +81-471-24-1501,
Fax:+81-471-25-1841)
FEATURES
source Location/Qualifiers
1..954
/organism="Mus sp."
/db_xref="taxon:10095"
30..527
/gene="mpin1"
30..527
/gene="mpin1"
/gene="mpin1"
/codon_start=1
/product="PIN1"
/protein_id="PAA87037.1"
/db_xref="GI:6468200"
/translation="MADEEKLPPGWEKRMSSKRSRYVYFNHITNASQWERPSSGSGTVG
GSSNNGGEPAPKVCISHLVYKQSRPSSRPQKIRSRKEALELNGYLOKIKSGEE
EDFESLASQSDSCSAKARQDGLGPFSGQMKQPEEDASFLARTGEMSGPVFTDSCGHIH
ILRE"
954
/note="30 a nucleotides"
BASE COUNT 209 a 260 c 295 g 190 t
ORIGIN

```

Query Match	6.2%	Score 61.2;	DB 94;	Length 954;
Best Local Similarity	50.6%	Pred. No. 0.0072;		
Matches 175; Conservative	0;	Mismatches 168;	Indels 3;	Gaps 1;

QY	444	tgaagatggtccaaagtgtagatgttcacatttctgtgttcaagaacaaatcaatcaagaanaacc	503
Db	185	tgAGGCTTCCAAGGtGGCGCTGCTCAATCTGCTGGTGAAGCAACACCAGTCTGGAGGCC	244
QY	504	caagctcttggaagtcctccagatggtataatgaactagaaactgagacgaactatcacagatatt	563
Db	245	CTCATCTCTGGCGG---CAGGAAAAGATCCACAGGACGAGGAGGAGGCCCTTGAACTCAT	301
QY	564	gaagaaacatttggaaagaataattttagtgtgtgtgttaaacctaagtgaaatttgagcaaatc	623
Db	302	CAATGGCTATATCCAGAAAGATTAACTCAGGAGAGGAAAGACTTGTGAATCTCTGGCTCACA	361
QY	624	cgaagatgatttcagctcacatlyacagaagtggtgtatbttagtggttllttagcaaaagaca	683
Db	362	GTTTCAGTGATTGGACGCTTGCCCAAGGCCAAGGGAGACCTGGTCCCTTCAGCAGAGGTCA	421
QY	684	aatgcacaccacattcgtgaagaagccgcattcaatttgcattgttggagaagtcaagtaaacat	743
Db	422	GATGCAGAAACCACTTTGAGAGTGCCCTTGTTGCTTACGAGGACAGGGGAGATGAGTGGGCC	481
QY	744	aattgaaacccaatagtgtgtgtccattactccaagaagaacagatata	789
Db	482	CGTGTTACGGAGCTCGGSCATCCATATCATCTCTGGCACAGATATA	527

RESULT	11
AC006201	
LOCUS	AC006201 87947 bp DNA PLN 05-APR-2000
DEFINITION	Arabidopsis thaliana chromosome II section 105 of 255 of the complete sequence. Sequence from clones T13L16, T27K22, F8D23.
ACCESSION	AC006201 AE002093
VERSION	AC006201.3 GI:6596825
KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryote: Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons, core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS	1 (bases 1 to 87947)
	Lin, X., Kaul, S., Rounsley, S.D., Shea, T.P., Benito, M.-I., Town, C.D., Fujii, C.T., Mason, T.M., Bowman, C.L., Barnstead, M.E., Feldblyum, T.V., Buell, C.R., Ketchum, K.A., Lee, J.J., Ronning, C.M., Koo, H., Moffat, K.S., Cronin, L.A., Shen, M., Vanhaken, S.E., Unayam, L., Talon, L.J., Gill, J.E., Adams, M.D., Carrera, A.J., Crasay, T.H., Goodman, H.M., Somerville, C.R., Copenhaver, G.P., Preuss, D., Nierman, W.C., White, O., Eisen, J.A., Salzberg, S.L., Fraser, C.M. and Venter, J.C.
TITLE	Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana
JOURNAL	Nature 402 (6763), 761-768 (1999)
MEDLINE	20083487
PUBMED	10617197
REFERENCE	2 (bases 1 to 87947)
AUTHORS	Lin, X.
TITLE	Direct Submission
JOURNAL	Submitted (09-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT	On Dec 17, 1999 this sequence version replaced gi:4406805. The sequence and annotation of chromosome 2 were merged from those overlaps. For detailed information, please see the TIGR web site (http://www.tigr.org/tdb/at/at.html).

Genes were identified by a combination of three methods: Gene prediction programs including GRAL (<http://arthur.epm.ornl.gov/pub/xgral>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://gnomic.stanford.edu/GENSCAN.html>), and NetPlantene

(<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on plant analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arjan Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/MashU/ABI consortium for sequencing BAC clones FBP23, F5J6, T17A5, and T13J16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jinting Jiang, Klaus Meyer, Eric Richards and Satoshi Rabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

FEATURES
source

1. 87947
Location/Qualifiers

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="11"

/complement(<1. 3627)

/note="Sequence from clone T13J16"

complement(join<343. .462,534. .743,840. .1391,1467. .1745,

1862. .2293,2413. .2841,2921. .3211,3306. .3566,3751. .5049,

5246. .5663,5754. .6166,6574. .6921,7017. .7274,7494. .8144,

8564. .8785,8983. .9204,9310. .9459,9927. .10317,10443. .10665,

10752. .11010,11311. .11646,11740. .12729,13148. .14380,

14459. .14485,14617. .14668,14760. .14823,15211. .15279,

15716. .15775,15858. .15920,16040. .16135,16394. .16502,

16675. .16760,17148. .17863,18061. .>18118))

/gene="At2g17930"

complement(<343. .>18118)

/gene="At2g17930"

/note="T13J16.5; similar to hypothetical proteins from S.

Pombe (C1F5.11C, PID:g1351684), S. cerevisiae (YHP9,

PID:g731689), and C. elegans (C47D12.1, PID:g1216120"

complement(join<343. .462,534. .743,840. .1391,1467. .1745,

1862. .2293,2413. .2841,2921. .3211,3306. .3566,3751. .5049,

5246. .5663,5754. .6166,6574. .6921,7017. .7274,7494. .8144,

8564. .8785,8983. .9204,9310. .9459,9927. .10317,10443. .10665,

10752. .11010,11311. .11646,11740. .12729,13148. .14380,

14459. .14485,14617. .14668,14760. .14823,15211. .15279,

15716. .15775,15858. .15920,16040. .16135,16394. .16502,

16675. .16760,17148. .17863,18061. .18118))

/gene="At2g17930"

/note="hypothetical protein"

/codon_start=1

/protein_id="A020114.2"

/db_xref="GI:6598526"

/translation="MSPIONFQHSRRLVDLPIPTRLMVEVRDSELIHAEVTL

NFLKYPASVLLQITKQFIDNPBHKRNIVELNRLPHSEVLRPVQDLKKA

MQVLADNENGLICRITFDLNRNPTLENEQPLDVCYISLFTVSHFEEN

VKKEEYKPMPTSDQSLTPPTPIGWOLNPTSTRKIIITSEPLVVMFLFOYSRL

QTNIPHLPLVAIVSGPENYPSHLKQFIEIKGQVKELVSLKHVGTGFEKRL

FPLIDILDERVIVGTGRACFESRLPLAYSLAEIVHVGDSLQSLSTILEFSN

MHSSTLSIHITTCARMLNLTVPDIEKGVDOOSMDARILLIGRIIDAFGKSTFER

TIQQLDEGEVGNDRYTLRSKLELFPVQVPEHSEKVENDCNLIKTLVGMCKITIMST

HAHLRFQGRKGRNDEVMKASGLKSLGVCHLALFKKDEKELNLFSSIIAIMEPR

DLMDNFSLCPPELPESMINNOUVOFPAALLQAPKRYKRPADVLINLVSSKLIDYKN

PDSAARGLVLHFRICIFGAVTKPSPFERLIQHNHVPINVECKNATEVEKPLGYQL

LRTFAVGLKXKYLLEIRLIPMLPLCKINLLITLMDGPGEDKDLIELCTLPARQL

SSLPLPLRMKPLVCELRGSDGLVSLGKLEFVNDLSNPDESLPMNVNSELIA

IMSHLRPVPMCKKALQILGKGRNRRKLEPLLECKNDPEHLRLVLPPEPTP

FLYLDKFTNLAVAIQNRHGDYIRKQALKEFLVCLLSOLNIEGCVTDNCGPRL

LSLTLSSVDSNMHSEAVEIKADLVKRTQIMARKSTFKLITLITLAA5DPDLS

TDDEFEENICRHAIIILHVDYSNASTSTSSLDG5VISTSSKSNOSNKLQDLP

IFPLDALVDYADNIRLHAKNALNALVWFATLLEFLARVHADLARGHNSMIVS

PSNPYSPHPSVRIVPEBOLPRLLHGCYSTWQMGVMGALGVKVNVELCY

FOYKIVRGVLYVLRKLPVYAKOEETSOVLILVNNVNDSEARSKRDQVE

YLATLEPMA3IIPVRKNONCIALASRPSQVTEPLTQIOLLOLIMPLRSEKT

VDQQVTAALNCLALRPLLVTPBELVNFDEALQIABEDTVAVNLMPKUTLS

LNRLFRACIELCTTMAWDFRTQTHNELKATISMFESLTCRAEIVAAVEGLRL

VINQOHPKELQSSLRPLIVNLAKHNSLMPLLLOGLARLDELATLVLIDLELA

HLKKVPEKELQSSOKAMKAGEEPKIAAIIIEFLHLLPHASKEFLDELATLVLIDLELA

LRPGVYSEINSPYRLPLTFKRLRYALAAVDEYLSHLSPEKYYRRMYIIR5AGQPL

RELASPOKTIISAPPELSPKPDPLSTTASAPRPTSSGDEHITVKLESSNAVTK

ANIASDVRQGLITLITMKKLIPSWLOSNSVDTLVILKSPARSIRQNBELMY

QVESKWLVCYFNLHHEKSEVNLLELISLTFHSRIDYTFLEKFIIEVABGTP

NMRALLDLHLNLFHSKQGLDHLVQAMOMLILPMLAHFONGQTEVEDIDPVKIV

ERLDEPEEYSAEDYDEPLREILQTLTLTKLYQSDLVHHRKELIFEGNNHLKREDA

SKOMAFVNCHEFLDAYOAPKILIOYFVALFTGCEPNKKLVQKALDILMPAPKRLP

LGSRMPDIWRTYRKTILVBERGSHIPLNHLIFLLVHRSDLFTYSCRAOPVPMYSR

LGLPYTTAENRRLAIELAGVYSNERQNRNKKMTDIDGTSQITDEHNTSGAPK

KRTDGSATSEDPKRYKIEPGLDSTCVMSGGASIPNVEPOSATOPDEERKPNAM

ESMIINFLRVALVIEPKRETYMTKQALDILSOLAEVPSANVFNVLKELSSMP

PSQSDSTALQGLDVMNKLVEKQPHLEIENNINOI5QILEPCFKHMKDAGSLSL

LKKWFAFPLDAANTPPDILKYOKNELINKNVNTVTAOTGSDN5GSI5FVLLV

IKTLAVNKNFVDSYLVLRRIORLQALDVSAGSHRQROGRTDSDSAVMSRQATD

AVTCNKSYLEIDETVYMLADCKRSYTOILLTLSEKGDASVILCTIDMKRWED

DFSKTASGLSGSF7QKDVLTLNKLSTIDKQHSSEALEEDQNTYDLOL7GLCD

TKYPLGLROVSLKVERHMLGRASHPGRRKFFLLYHELSLKTFLARLOYTIOJD

WEALSDVFMQLQGLDLALLVLEDEKISLAPNARFPLPSNP5G10QAPNLEGP

EEVTSMPDSIVMKHAOELSATSKLOVADVTPRELHADVANA4YMLVPEIYVWT

LKEBOVALKPMISILSKDYHKKQOGRPNVVOALLLEGOLSHQPRMSELIKXIG

KTYNANHLALLESVYMLPMDNSKCAESAEILRLNEDNDRFGIMKRSTIETRA

GLSVLQHGFEWQROASLFTQAMVAVATQGTNNVTPAKEMCLMEQMLHCSQ5QWAL

VDSKSTENVEILDLISLMKLPDVAYLKDIH1PAQVVEPKRLVQSYFALHDNRNG

VGAEVNTGKGVGLAEOMQOLPEMSVNAHVP1LOOFOOLVEOESAR1HVDNRNG

VSGNTPVGGVGNAYDLKDLIETWRLRTPREMNMMVWYDMLQMRMNNV1DARKD

FATNSNPRIHIGFRDKAMVNYKTLARROKIDVQVOLLERKYGSTMEGSAFKI

REOAKYVLEMKGRASGLN1JNSTNLETPEDTKAKELFRLEKDFHLKNDSEANAY

SNATITFLKMLPKMISWGSYCDAAVYQDEDEWLEVA5FCLOGIFRGVSN5SHAR

VLYLVLFYDTEINPEVGRVFKHLDQVPHWML5WIPOLL5LQTEAPHOKLYLKTAA

VFPQALYMLRTYLLERRDQVNVKSELRVLVLAQRMQONATGAGHGSN1PSENOIHQ

AQTSAGGATDSCGNPHGDOESERTENNLLPVGSDQMHQSS5L1INNENETVARNAS

LALSAGCAPDAADIMBALRGKNNNLASLVEVLLTFEGSRFVLLPREELAVVNLHL

RCTKTPATTAEVPOPLKEKELSGVCACRSADAVTHAVEYKTKODFERHDPSTT

TEPATYATVELARLKMKNNLIQSNVEDRFPVVLLEDSQKRLRFNVNDE1PQVYAD

QEVAPHTVLELDVGVADP1VRNHDG5FRRLTLDGDSQKRLIYQTS1LTPNARSDR

ILDLFRMNQMPKHKESRRRHIGITPIIIPW5QVNRWEDMLNTELEVYENCA

RNLRADPLRTHREKOLNOAISQISAEALRDAIYDITLTVNDS1F5QYVNTK

LMGSHMNAKKQFAVOIAYVSSPM5M2LOIGSGSPKVLPAKTKGMPOTDRPADA

NGMIEHREYPERFLTRNMQAF5QFEVBEGLSKMSKSAQAV1SSQNHRLRQALM

FRQELLSWFRRLGVPIIPVGGIATLNPALKKNANVEDYIKRIRGIAQYF5EE

DENTVEP5QVQGVLEWLEALSPNNL1CMQDPTWHPW"

complement(3628. .86350)

/note="Sequence from clone T27K22"

11013. .11066

/rpt_family="Poly_A"

/join(<18766. .18903,19215. .>19703)

/gene="At2g17940"

<18766. .>19703

/gene="At2g17940"

/note="T27K22.19; predicted by genscan"

join(18766. .18903,19215. .19703)

/gene="At2g17940"

/note="hypothetical protein"

/codon_start=1

/protein_id="A020132.1"

/db_xref="GI:4406824"

/translation="MERGNWNSGRABDIETKAAFGSVKAEVAMFGEKVLAGEIYATRL

Query Match	Best Local Similarity	Score	DB 14	Length	1023
Matches 101, Conservative	5.7%;	57.7%;	Pred. No. 0.062;	Mismatches 74;	Indels 0; Gaps 0;
Db	614	tgccaataaccgaaagtattgcagctcatcatgacgaaggtggtgattagagttttaa	673		
Db	488	tgccgtaaaagagacagatcgttccttcatatcmaacgagggcgacccctgccttcg	429		
Db	674	gcaaaagacaaatgacacaccatcgcgaagaagccgacatcaattgcatgattgagaag	733		
Db	428	ggagagggcagatgcagccttagcttgaagacgctgccttccagctcaaggctggcgag	369		
Db	734	tcaagataaattgaacacaaatagtggtgctcatatcctccaagaacagata	788		
Db	368	taagcagatagctgtaatcaggaacgggtgctcatgcatcaacggcgtaggtta	314		
RESULT 14					
LOCUS	SCPTFL	2689 bp	DNA	PLN	17-FEB-1997
DEFINITION	S. cerevisiae Pflp gene.				
ACCESSION	X85972				
VERSION	X85972.1	GI:758283			
KEYWORDS	Processing/Termination Factor 1; Pflp gene.				
SOURCE	baker's yeast.				
ORGANISM	Saccharomyces cerevisiae				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;				
AUTHORS	Saccharomycetaceae; Saccharomyces.				
TITLE	1 (bases 1 to 2689)				
JOURNAL	Engelbrecht, J.A., Voelkel-Melman, K. and Roeder, G.S.				
MEDLINE	Meiosis-specific RNA splicing in yeast				
REFERENCE	Cell 66 (6), 1257-1268 (1991)				
AUTHORS	92005676				
TITLE	2 (bases 1 to 2689)				
JOURNAL	Hani, J., Stumpf, G. and Domdey, H.				
MEDLINE	Pfl1 encodes an essential protein in Saccharomyces cerevisiae,				
REFERENCE	which shows strong homology with a new putative family of PPIases				
AUTHORS	FEBS Lett. 365 (2-3), 198-202 (1995)				
TITLE	3 (bases 1 to 2689)				
JOURNAL	Hani, J.				
MEDLINE	Direct Submission				
REFERENCE	Submitted (28-MAR-1995) J. Hani, Genzentrum der Universitaet				
AUTHORS	Muenchen, Muenchstrasse 221, D-81375 Muenchen, FRG				
TITLE	4 (bases 1 to 2689)				
JOURNAL	Hani, J.				
MEDLINE	Direct Submission				
REFERENCE	Submitted (13-JUN-1995) J. Hani, Genzentrum der Universitaet				
AUTHORS	Muenchen, Muenchstrasse 221, D-81375 Muenchen, FRG				
TITLE	Related sequences: M38340 and p22696.				
JOURNAL	location/Qualifiers				
MEDLINE	1..2689				
REFERENCE	/organism="Saccharomyces cerevisiae"				
AUTHORS	/strain="DH484"				
TITLE	/db_xref="taxon:4932"				
JOURNAL	/clone_lib="A. Haid"				
MEDLINE	/clone="1"				
REFERENCE	1..310				
AUTHORS	/gene="WER2"				
TITLE	<1..310				

Query Match	Best Local Similarity	57.7%	Score 56.6	DB 14	Length 2689
Matches 101	Conservative	0	Mismatches 74	Indels 0	Gaps 0
614	tggaataacgcaagaatgattgcagctacactagcaagaagtgatattaggtttta	673			
2169	tgagctaaagagagatcagattgctcttcatatcagcagagcgacacctcgcttgcc	2228			
674	gcaaaagacaatgcaaacaccatcgaagaagccgacatcgaattgcatgttggaagaag	733			
2229	ggagagggagagtgagagcttgagcttggaagacgctgcttcaagctcaaggtcgagcagag	2288			
734	tcagtaacataatgaaacaaatagtggtgtccatatactccccaagaacaggata	788			
2289	taagcgatattcgttgaatcaggaacgctgttcatgatgatcaagcggttaggtta	2343			

RESULT 15	SCXCSOM83	43661 bp	DNA	PLN	11-NOV-1996
LOCUS	S.cerevisiae chromosome X DNA (cosmid 83).				
DEFINITION					
ACCESSION	X87611				
VERSION	X87611.1	GI:854567			
KEYWORDS	adenylate cyclase; alpha-eggultulin; ATP sulphurylase; beta-adaplin; CEN10 region; CoA thioesterase; CYR1 gene; dilydroxyacid dehydratase; ess1 gene; glyceraldehyde-3-phosphate dehydrogenase; ILV3 gene; mer2 gene; MET3 gene; su12 gene; TDH2 gene.				
SOURCE	baker's yeast.				
ORGANISM	Saccharomyces cerevisiae				
REFERENCE	Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
AUTHORS	1 (bases 1 to 43661)				
TITLE	de Haan, M.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (22-MAY-1995) M. De Haan, University of Amsterdam, Dept. Molecular Biology, Kruijslaan 318, NL-1098 SM Amsterdam, NETHERLANDS				
AUTHORS	2 (bases 1 to 43661)				
TITLE	Fang, H., Panzer, S., Mullins, C., Hartmann, E. and Green, N.				
JOURNAL	The homologue of mammalian Spc12 is important for efficient signal peptidase activity in Saccharomyces cerevisiae.				
MEDLINE	J. Biol. Chem. 271 (28), 16460-16465 (1996)				
FEATURES	96279206				
SOURCE	Location/Qualifiers				
gene	1. 43661				
CDS	/organism="Saccharomyces cerevisiae"				
	/strain="FY1679"				
	/sub_strain="5288C"				
	/db_xref="taxon:4932"				
	/chromosome="X"				
	/clone="cosmid83"				
	1. 4065				
	/gene="CYR1"				
	<1. 4065				
	/gene="CYR1"				
	/codon_start=1				
	/product="adenylate cyclase"				
	/protein_id="CA60917.1"				
	/db_xref="GI:854568"				
	/db_xref="SWISS-PROT:P08678"				
	/translation="DPRHYAIRFENDDPTFTLSCPPATVEEIIIPALKRFPNTACGNPISLKVGLSKLIRPTSKLIERKLLINCRSDPLHIGEDISLSPFEPFPPVPSHSHPPEBOEORIMSEPEHYVLDLRNDDLTTPPIITQHTSEIESLDVSNANITFLELFESSIKILSLRMVNI RASRPPSNITAKRVLSELONNR IKRVPSIMKLSLTLT LNCOCNELESIPAGVELKLNQILDSSNKMHVPEVNYCTNLDSYKLOSLEOSTYLVAKMANSKLNKLTIDGLESTDELTPTLNKRNRSLIKTANSNOMNFETNRISEFDTLPKRLALEIOENPRTISGPDKOENPMNLSITLNKOLSSIPBELTILKLS PLELELNQNNLTRLPDEISKLTVLVLAVSARAKLEIYTPPLSOLSKSLRTIDLSNNIN RDPFDGEMNELTSLNITSSNAFGSSLENSYVHNMSYGSLSLMEFTADNOFDDM MPPLFNCVNLKVLNLTSSNNPDSVSHMKLESITELYLSGNLTTLTSGDTVLKWSLKL LMLNSQMLSLPAELSLNSQLSVDDVANGOLKYLISNHYDMNRNKKELLYLNEGSKL RREFIKSFIDHDIDALDVLTPOLKYLGLDMVTLNMTKYPDENVNRLRTTASIIIN GMRVGVAADTLGORDYVSSRDYTERFERGNDCECLLHDSKNQADGHNISRIYRDL YDKLILQLEERTGDETDNDIKTALRFSTLQNKTINGLNSVNDGADVANLSTVDLSL GACSTVITYIRKRLKFLAANLGDCAMLLSKNNGDYOTLTKOHLPTKREVEYIRISGGVY L NNGTLDGVDVDSRAVGFEDLLPHTHASPDLSVNLTRTADEMLIVATIKHLYEMDVDTYVTCIDIRKNSDTPDLRAAELKDHMAVYGCENTITICLALYENIIOONRPTLTKNSMLTMR RSTPEEDTIRLROPEISIPPTGILAMVETDISKSFMEELPNMARTAKTHNDIMRRGLRIYGYEVKTEGDAFNAPFTPTSGLTWCLSVOLKLLDAOWPEEITISYDGCYOTDB NGNITLYGLSVRMGIHMGCPVPELDLVTYQRMDYIGPMVKNARVGAADGGQILAMSSSIFYSFNKIMKHYHVRVNGKESLKEVGEIIEVLERIAMESIGNAFPDFGEHKL L GLEKRELVITIAPEKILASRHEFASEDEOSKLIINFTMLFRLVINSALSGKLEIIDSRTSGSYIKFNPVENCIMQISISKDALLFEDHYITRIESSVALHLHQRCSS LEICRDNKTSARSNFVNDLDMVKNARLST"				
gene	complement(4417..5028)				
	/gene="ORF YJ783.17"				
CDS	complement(4417..5028)				

gene
CDS
/gene="ORF_YJR83.17"
/codon_start=1
/protein_id="CA60918.1"
/db_xref="GI:854569"
/db_xref="SWISS-PROT:P41544"
/translation="MAVSTRKRYRNPENLKPQIQIEKQDLSPSKIGIQIVLQIFVYVTT
AKLVLPYCKAKLGYDLNLIKEMLEFSENIDFPNAGISLSMLKGLCVLFIYVYQ
RKLAMDVAITTHAINFIVLFTYRKPSPFSWFFQIILSLILFLPLGTTTRMRELRL
TFEEELIYVNPGEVGLVTPSQOHSNHSLEQSPIDQLKLEQSI"
5472. 5828
/gene="ORF_YJR83.2"
5472. 5828
/gene="ORF_YJR83.2"
/codon_start=1
/protein_id="CA60919.1"
/db_xref="GI:854570"
/db_xref="SWISS-PROT:P47081"
/translation="MSESGKKRSPROOLYEAISLAGYKKALSNHPEFGLPEPCAT
IYLGSPWSSFPALKEQGRKQVEINDEILIKTKNOREDFIDKEEYIRLOGLSBEDRM
EYRVARLKDESENW"
/gene="RIB1"
complement(6049. 7479)
/gene="RIB1"
complement(6049. 7479)
/gene="RIB1"
/note="homologue of the human /rat ribophorin I precursor
protein RIB1"
/codon_start=1
/protein_id="CA60920.1"
/db_xref="GI:854571"
/db_xref="SWISS-PROT:P41543"
/translation="MKQVFSWIVGLFLCFPNVSSAOYEPPATWENVYKRTIDVSN
AIVSETIKINIASEPATEYFPAFEGSIFSKVSFSAYFTNTEATFLNSQLANSTTT
AGDEDSKIRYGIIOFPNALISQDEVLVLSKFPNTVGIYPEHAWMSEDOHLMEET
NGLPSAYDTKASFTLLGSSFEHYPHNESDLKANGNSFEEGPMEDIPRSNSNE
TLAIYVSHNAPLVQVNLBRDILWHSMASTIOFEERYLTLLTKAKALSFGFRLMLKQ
IOTQNMROTHFATVILDMILPEGATPHYETDVLGAYTSHARDHFIETPRPITGGCA
YNFTVGMTHKLSDFLHWSGSDKEFYVASIPILNGPPDVIYDVELSVFLPECAELFDI
LSIAPYFIALSVETQKSYFDPLNGHVKLTFESYRNLSIQVANQVLIKIDYPRSSFFKKR
LSPVYIFPILMGVFLVLTLMNMVTN"
8173. 8754
/gene="proteasome component pre3"
8173. 8754
/gene="proteasome component pre3"
/codon_start=1
/protein_id="CA60921.1"
/db_xref="GI:854572"
/db_xref="SWISS-PROT:P38624"
/translation="MAVYPRKQVYLIGASRITTCAYINRYTDKLRVHDKINCRSGS
SAADYQALNDIVQVLELITSOYGTGPSTETASVKELCYENKMDLTAGIIVAGTDD
NGEYVITPDLGGSVHKLPYALVLAGSGSTETIYGCDKNFRENMSKEETVDFIKHSQAI
KWDGSSGVIRMWVTLAAGVERLIFYPEVEQL"
complement(9162. 9223)
/note="CEN10 (centromeric region)"
9630. 11438
/gene="ORF_YJR83.4"
9630. 11438
/gene="ORF_YJR83.4"
/codon_start=1
/protein_id="CA60922.1"
/db_xref="GI:854573"
/db_xref="SWISS-PROT:P47082"
/translation="MPEQELSPNGRKSEVHYIISIPLNRSAPSPDDSVSQFOSDFE
MYRQSIIDHPGSPFGVNSLSPKTSIRANSFENITELANDNERSFPEKSENDEYDVE
DTLPAALDGRRLSVTLNNAGRPRITNLANDNRVPSASMAIIDDYGSIONSTIGDSS
ILTRPASIEMKSGAGRRFTNDDMSITVYKRVESVDCKVYTLLAGOSTAQTIFNSI
NVLIIGLIGLAPLGKIVAGWVIGLTMALIFALFFCAVELISCTLDQDPTLISVADGL
YAAFTSGKALISALFTLDLIGSGVSLVILFGDSSLALFPQYSTFFFIYVFVTPPE
VETIPSVYSNLSILGLISTGTGVIVICCGGYKSSPGSLVNPMTSMWPIDLKHLLC
IDELISACMGGAIVPEPNLKTDRMBHDKCDKLKTTKYTSTVDIGTAVIGLHGNLIV
KDEITKNVLTREGEPRKYVYGLISALMTIIPILAKRPIVSVIVDLMVNHQIHDEAA
SAIKRRAKGLDVEGRNRTINVFVLIATNFEPEFDITLNFAGLCFTTICLLTPCWFTVL
RLCKTITIPMERVAVCHVIGICISVLSLTIGVAAIIS"
11692. 13473

```
CDS
/gene="ORF YJR83.5"
11692..13473
/feature="ORF YJR83.5"
/codon_start=1
/db_xref="GI:854574"
/db_xref="GI:854574"
/translation="MSLFGVLKSNAGRIILKDPSSATSKDVAKYIDSVINTCKNGSIT
KRALEDEITYGLDANQWOMVKLVLSIDGLIOGLOELKDVYTPSHNSLSDGSLNS
SGSESELEPEASVEPKKQMLSDVSEIEQSDSLSENDEPSPMDKETSARARE
FAEKRISSGQDERHSSPDYGINDKFTDLKFNROT LAEDSNEASEGSEDIDTF
QDMPSDDEEEAIIYEDFDKPKPEPVKHSVDKPERDEDELEEDHSAMDVKYLDL
FADBEDEPNAGEAEDSKNLSSEKQOIEIRKOIEQLENEAVERKMSLKGEVAKD
RPEDALITELEEDRTAKPVPTTSEVTESELEDMIRRIODSNEDLORRTLDITRK
SORPELSDVKSSELAIEYEDDYTRADESEALSELOKASEISELYANLYVKKLDY
LSVHFVPKPAASLEIRVETPTISMEDAOPIYMSNASSLAPOEIYNVCAEKDEGR
LKNGVAMSKKEELTRDEKRNRLRALKRKSKANLPVYKRSKRDVVDTLKAKNITVI
NOKGRKDVSGKTKRSRSGPDSINIKL"
complement(13821..15440)
/gene="ORF YJR83.30"
complement(13821..15440)
/gene="ORF YJR83.30"
/codon_start=1
/protein_id="CAA60924.1"
/db_xref="GI:854575"
/db_xref="GI:854575"
/translation="MKFDESYFFKCLTANRHSIMLSLHSAATLSNKRFPYLSHSN
RKNIILKLRHPSFDPIRHHLPEDITTDIYPSISQNVYESLANKLEVPKDAAMVHNM
IENLSDIDYGVATIHSSNLRDLKSLPAIKOIRNNRGRVSSMELPTQYKSMEN
VPELMEVYIEKLIKFDKAEKVDGKSLTODLVRCILYLNHPSNINLSELVEPLT
IYVDNGIPVNLGSLVKRKIPLSFEDKYVSEMYOQICELDYFSLDNIYADPLVLEK
CLTVLENEKIQOTEKEKELISKLEEDIDYKSOCHDNMSLEFPNMSVRKATSFEL
FLEIQKRNIDKDFELAKLRLIGAFKGVSLFKLYDEYLLFKNNEDDLMEFAEL
TLCOCQKSSNEKLOYAEAFIKEDFDSKESIOSVLIVANAKANIDSLIKYNSNI
STAKRKDKYTDIAESDVLTESLILAFISRDADPARVIFDGAIGELISGPTAAKKI
KNLLAOGALERTKSKQVMQTKIEHYMESI"
complement(15736..17688)
/gene="AGAL1"
complement(15736..17688)
/gene="AGAL1"
/codon_start=1
/product="alpha agglutinin"
/protein_id="CAA60926.1"
/db_xref="GI:854577"
/db_xref="GI:854577"
/db_xref="SWISS-PROT:P20840"
/translation="MFTFLKILMLFSLALASAININDIYFSLMETPTLPANKPODG
WTFPFESIADASSIREGDEFTLMPHYRIKLNSOTATISLADGTEAFKCYVSOQ
AALVETNETCTCAONDISYNTIDGSIPTSLNPSDGSSEYELENAKFRKGPMLY
KLGOMSDVYNDPDAFTENVRHSGSTGTGFSYHLGMCNGLFLGSTERIDYDS
SNMNVLDLSSVQVYSSNDPDMWFPOSYNDYNADVCFGSNMLITLDEKLYDEMILM
VNALQSLPANVNTIDHLEFOYCLDTIANITYATQVSTTREFIVYQGNMLGTASAKS
SFSTTTDLTSLINTSAYSTGSISETVETGNRTSEVISHVVTSTKLSPATTSITIA
QNTSYTSDNITVGTIDIHSTSEVISDVERTASTRETAIVAAVPTSTKSGAMNTYISO
FTSSSPATINSPILISSAVERISDASIVNHTENTNTRAAPSEPTFYNATRNLSLN
SPSSKQPSPSSTYSSPLVSLSVSKTLTSLTSPSPVPSNTYIKTKNTGYEHTAL
TTSVGLNFSFETAVSSQGTIKIDTFLVSSLIAYSSASGSLG100NFTSLMIST
YEGKASIFFSAEIGSIIFLLISYLF"
complement(16827..17150)

gene
Query Match 5.7%; Score 56.6; DB 14; Length 43661;
Best Local Similarity 57.7%; Pred. No. 0.045;
Matches 101; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 614 ttgcaaataccgaagtgatgcagctcacatgcagagtggtgatttaggttttta 673
|||||  |||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 39525 TGGCTAAAGAGATCAGATTGCTCTTCATACAAAGCAGGCGGCGACCTCGGCTGGTTCCG 39466

QY 674 gcaaaagacaatgcacaaccacatgcagaagcgcatcoaatlgtcatgttggagaag 733
|||||  |||||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 39465 GGAAGCGCGAGATCGAGCTTGAAGACGCTCGCTCAAGGTGCGGAGG 39406

QY 734 tcagtaacataatgaaccaatagtggttcacataatccccaagaacagagata 788
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

Db 39405 TAAGCATATCGTTGAATCAGAGACGCGTTCATGTGATCAACGGGGTAGGTTA 39351

Search completed: June 8, 2001, 06:02:12
Job time: 7262 sec

PT Novel Candida albicans gene, CaESS1 useful for identifying compounds
PT that specifically bind to and/or inhibit CaESS1 and thus for treating
PT Candida albicans infections and other life-threatening fungal
PT infections -
XX
PS Claim 2; Fig 1A; 51pp; English.
XX
CC This is the complete nucleotide sequence of the CaESS1 gene from
CC Candida albicans. The gene encodes a 177-amino acid protein,
CC CaESS1 (see Y95876), which is the C. albicans homologue of
CC Saccharomyces cerevisiae ESS1 (42% amino acid identity). The
CC CaESS1 gene was isolated from a C. albicans genomic DNA library by
CC functional complementation of a temperature-sensitive S. cerevisiae
CC strain, ess1-L94pts. CaESS1 nucleic acids, especially CaESS1-specific
CC primers (see A50293-94) and probes to determine the presence of C.
CC albicans in a sample or specimen. CaESS1 protein is a target
CC for screening for antifungal and inhibitor compounds, useful for
CC treating or preventing C. albicans infections. CaESS1 DNA can also
CC be used to generate diagnostic probes or primers for replicating or
CC cloning C. albicans DNA. A CaESS1 inhibitor can be a compound
CC which selectively inhibits growth of S. cerevisiae not containing an
CC endogenous ESS1 gene but rather CaESS1 and uninduced PIN1 (a human
CC e.g. does not inhibit S. cerevisiae and/or preferably does not inhibit induced PIN1,
CC ESS1 but rather induced PIN1. Compositions which inhibit PIN1
CC are useful antiproliferatives e.g. anti-neoplastics, antitumour
CC agents or anticancer agents.
XX
SQ Sequence 989 BP; 340 A; 174 C; 190 G; 285 T; 0 other;

Query Match 98.5%; Score 974.6; DB 21; Length 989;
Best Local Similarity 99.1%; Pred. No. 2.8e-236;
Matches 980; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 gataaccaatagatgtgtgtgtcctaaccagaagcggcttgaagaagaggtatt 60
DB 1 gataaccaatagatgtgtgtcctaaccagaagcggcttgaagaagaggtatt 60
QY 61 tagacacaagaagcattagcacttggaatagatatacaagttgagcttccttcgataga 120
DB 61 tagacacaagaagcattagcacttggaatagatatacaagttgagcttccttcgataga 120
QY 121 tattaagtagtgatattacattacaacaaactctctcttcttataattctatcaacaaca 180
DB 121 tattaagtagtgatattacattacaacaaactctctcttcttataattctatcaacaaca 180
QY 121 tattaagtagtgatattacattacaacaaactctctcttcttataattctatcaacaaca 180
DB 121 tattaagtagtgatattacattacaacaaactctctcttcttataattctatcaacaaca 180
QY 181 agatttcgtgtgcttctgtgtatatttgcacagtttagcttgattcttctt 240
DB 181 agatttcgtgtgcttctgtgtatatttgcacagtttagcttgattcttctt 240
QY 181 agatttcgtgtgcttctgtgtatatttgcacagtttagcttgattcttctt 240
DB 181 agatttcgtgtgcttctgtgtatatttgcacagtttagcttgattcttctt 240
QY 241 cagtagtatatcatatgcatgcacatcaacagcgtcttaccacaaatttgacgattaga 300
DB 241 cagtagtatatcatatgcatgcacatcaacagcgtcttaccacaaatttgacgattaga 300
QY 301 gtatccagatccataaagaagatattctttaaccaatctacacaaatgagtcgtcttg 360
DB 301 gtatccagatccataaagaagatattctttaaccaatctacacaaatgagtcgtcttg 360
QY 301 gtatccagatccataaagaagatattctttaaccaatctacacaaatgagtcgtcttg 360
DB 301 gtatccagatccataaagaagatattctttaaccaatctacacaaatgagtcgtcttg 360
QY 361 gaccacacttctgcatcgtgacaagaagatattgaaatgcatctgcgaattttaaanaac 420
DB 361 gaccacacttctgcatcgtgacaagaagatattgaaatgcatctgcgaattttaaanaac 420
QY 421 aatgttcaagcagcactgtgaatgagatggcaggttagaatttcattcttcttgatc 480
DB 421 aatgttcaagcagcactgtgaatgagatggcaggttagaatttcattcttcttgatc 480
QY 481 aagaacaatcaatcaagaanaacccaagctcttgaagtcctccagatggtatagaact 540
DB 481 aagaacaatcaatcaagaanaacccaagctcttgaagtcctccagatggtatagaact 540
QY 541 agaaacgagatctatacagatatttgaagaacatttggaanaaatatgagtggtgaggtt 600
DB 541 agaaacgagatctatacagatatttgaagaacatttggaanaaatatgagtggtgaggtt 600

DB 541 agaaacgagatctatacagatatttgaagaacatttggaanaaatatgagtggtgaggtt 600
QY 601 aaactaagtgaattgycacaataccgaaagtgtatgcagctcacatgcagaagtgatg 660
DB 601 aaactaagtgaattgycacaataccgaaagtgtatgcagctcacatgcagaagtgatg 660
QY 661 ttaaggttttttaagcaaaagacaatgcaaccacattcgaaagaagccgcatcaatttg 720
DB 661 ttaaggttttttaagcaaaagacaatgcaaccacattcgaaagaagccgcatcaatttg 720
QY 721 catgttgagaagtcagtaacataatattgaaaccaaattggtgtccatctccaaaga 780
DB 721 catgttgagaagtcagtaacataatattgaaaccaaattggtgtccatctccaaaga 780
QY 781 acagataatcaagatattgagttgatatataatgaaataatagacaaagtgtg 840
DB 781 acagataatcaagatattgagttgatatataatgaaataatagacaaagtgtg 840
QY 841 atagatttgtaacccaanaaagcagatgctcacaanaagtcgaaaactgtgagaagaat 900
DB 841 atagatttgtaacccaanaaagcagatgctcacaanaagtcgaaaactgtgagaagaat 900
QY 901 cttaacaggtacacggcgatataaacttaactgctgatatattatataatcgagaagtt 960
DB 901 cttaacaggtacacggcgatataaacttaactgctgatatattatataatcgagaagtt 960
QY 961 ccgcgcatgtgttgtatatttgatcc 989
DB 961 ccgcgcatgtgttgtatatttgatcc 989

RESULT 2
C33551
ID C33551 standard; DNA; 747 BP.
XX
AC C33551:
XX
DT 17-OCN-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 3467.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
EP1033405-A2.
PF 25-FEB-2000; 2000EP-0301439.
XX
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140693.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142053.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144326.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145917.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145919.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.

PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	22-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161922.
PR	28-OCT-1999;	99US-0161935.
PR	29-OCT-1999;	99US-0162142.
Query Match 8.8%; Score 87.2; DB 21; Length 747;		
Best Local Similarity 5.9%; Pred. No. 8.2e-13;		
Matches 207; Conservative 0; Mismatches 168; Indels 9; Gaps 1.		
OY	415 aaaaacaatggttacaagaaccacttgtaatsgagatgsgccaggtttagatttcacattg	474
Dd	159 aacaagcatcgaaagaagagcagagtgcgtcgtagagagacgaagttaaagcatalccacatt	218
OY	475 ttgatcaagaacaatcaatcaaagaanaaccgaagctcttggaagttcccagatgg-----	527
Dd	219 tfgattaagcatcacagggtcttaggsagaaagcgctcgtygaagatccagaagsgaagatt	278
OY	528 --tataaagtagacacagacgaatctatacagatatgtgaagaanaacatttgsaagaata	585
Dd	279 attctgcatcaccacatcagaagaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	338
OY	586 ttgaatggttgaggtttaactaagtlbaatlgycaaatlacccaagtgatctgcagccacat	645
Dd	339 gctcgcgcaagsgcgaatttcgaagaagtgysgactcgtgttcctgcagctgtacctgct	398
OY	646 gacgagaggtgtgtgatttgagggttttttagcnaaggaacaatlytgaaccaccalcgaaaga	705
Dd	399 aaacgc	458
OY	706 gccgcattcaattgcatglttgagagaagtcagtaaacataatgtgaacccaatagtggtgc	765
Dd	459 gaacctaagcaccatcaaggttggagatataaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	518
OY	766 catactccccaagaagacagataa 789	
Dd	519 cacatcattaaagagacagcttaa 542	
RESULT	3	
ID	C47563	
XX	C47563 standard; DNA: 833 BP.	
XX	C47563:	
DT	18-OCT-2000 (first entry)	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 54285.	
KW	Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.	
OS	Arabidopsis thaliana.	
PN	EP1033405-A2.	
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
PR	25-FEB-1999; 99US-0121825. 05-MAR-1999; 99US-0123180. 09-MAR-1999; 99US-0123548.	

PR	23-MAR-1999;	9905-0125788;
PR	23-MAR-1999;	9905-0125786;
PR	25-MAR-1999;	9905-0126254;
PR	29-MAR-1999;	9905-0126785;
PR	01-APR-1999;	9905-0127462;
PR	06-APR-1999;	9905-0128234;
PR	08-APR-1999;	9905-0128714;
PR	16-APR-1999;	9905-0129845;
PR	19-APR-1999;	9905-0130077;
PR	21-APR-1999;	9905-0130449;
PR	23-APR-1999;	9905-0130510;
PR	23-APR-1999;	9905-0130891;
PR	30-APR-1999;	9905-0131048;
PR	30-APR-1999;	9905-0132048;
PR	04-MAY-1999;	9905-0132407;
PR	05-MAY-1999;	9905-0132486;
PR	06-MAY-1999;	9905-0132487;
PR	07-MAY-1999;	9905-0132863;
PR	11-MAY-1999;	9905-0134256;
PR	14-MAY-1999;	9905-0134219;
PR	14-MAY-1999;	9905-0134221;
PR	14-MAY-1999;	9905-0134370;
PR	18-MAY-1999;	9905-0134376;
PR	19-MAY-1999;	9905-0135194;
PR	20-MAY-1999;	9905-0135421;
PR	21-MAY-1999;	9905-0135353;
PR	24-MAY-1999;	9905-0135629;
PR	25-MAY-1999;	9905-0136021;
PR	27-MAY-1999;	9905-0136392;
PR	28-MAY-1999;	9905-0136782;
PR	01-JUN-1999;	9905-0137222;
PR	03-JUN-1999;	9905-0137528;
PR	04-JUN-1999;	9905-0137502;
PR	07-JUN-1999;	9905-0138074;
PR	08-JUN-1999;	9905-0138094;
PR	10-JUN-1999;	9905-0138540;
PR	10-JUN-1999;	9905-0138847;
PR	14-JUN-1999;	9905-0139311;
PR	16-JUN-1999;	9905-0139452;
PR	16-JUN-1999;	9905-0139453;
PR	17-JUN-1999;	9905-0139492;
PR	18-JUN-1999;	9905-0139454;
PR	18-JUN-1999;	9905-0139461;
PR	18-JUN-1999;	9905-0139462;
PR	18-JUN-1999;	9905-0139463;
PR	18-JUN-1999;	9905-0139457;
PR	18-JUN-1999;	9905-0139458;
PR	18-JUN-1999;	9905-0139459;
PR	18-JUN-1999;	9905-0139460;
PR	21-JUN-1999;	9905-0139817;
PR	22-JUN-1999;	9905-0139899;
PR	23-JUN-1999;	9905-0140353;
PR	23-JUN-1999;	9905-0140354;
PR	24-JUN-1999;	9905-0140693;
PR	28-JUN-1999;	9905-0140823;
PR	29-JUN-1999;	9905-0140991;
PR	30-JUN-1999;	9905-0141287;
PR	01-JUL-1999;	9905-0141842;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142055;
PR	06-JUL-1999;	9905-0142390;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142920;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143542;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;

PR	16-JUL-1999;	9905-0144086;
PR	16-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	19-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144884;
PR	21-JUL-1999;	9905-0144814;
PR	21-JUL-1999;	9905-0145086;
PR	21-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145115;
PR	23-JUL-1999;	9905-0145142;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145324;
PR	26-JUL-1999;	9905-0147260;
PR	27-JUL-1999;	9905-0145913;
PR	27-JUL-1999;	9905-0145918;
PR	27-JUL-1999;	9905-0145919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	02-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0147034;
PR	04-AUG-1999;	9905-0147208;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148655;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149368;
PR	17-AUG-1999;	9905-0149175;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	20-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	25-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151343;
PR	01-SEP-1999;	9905-0151509;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153758;
PR	15-SEP-1999;	9905-0154018;
PR	16-SEP-1999;	9905-0154039;
PR	22-SEP-1999;	9905-0154779;
PR	22-SEP-1999;	9905-0155139;
PR	23-SEP-1999;	9905-0155486;
PR	24-SEP-1999;	9905-0155569;
PR	28-SEP-1999;	9905-0156458;
PR	29-SEP-1999;	9905-0156545;
PR	04-OCT-1999;	9905-0157117;
PR	05-OCT-1999;	9905-0157753;
PR	06-OCT-1999;	9905-0157865;

PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158368.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

[illegible]

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14609.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

```

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155466.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159328.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match Best Local Similarity 8.8%; Score 87.2; DB 21; Length 835;

Matches 20; Conservative 0; Mismatches 168; Indels 9; Gaps 1;

```

QY 415 aaaaaaattggttaaacgacattgtgaatgagatgagccaggttagatttcatttg 474
DB 145 aacaagcctcgaagaagagacagatgctgcgagagaccaaagtttaagcattc 204
QY 475 ttgtatcaagaacatcaatcaagaacaccaaagttcttgaaagttcccccagatgg 527
DB 205 ttgtatcaagaacatcaatcaagaacaccaaagttcttgaaagttcccccagatgg 527
QY 528 --tataaataaactaagacgaatctatacagatatgaagaacatttgaaagaata 585
DB 265 attctgaactaccactaaggaagccgcgtcgagcagcttaaatcgatccgtgaagatatt 324
QY 586 ttgagtggtgaggttaaacatgaatgtgcaataacagaagtgaatgcagctcacat 645
DB 325 gtcccgagcaagcggaatttcgaagaagtgcgactcgtgttctgactgtagctcgt 384
QY 646 gaacagagtggtggttttagggtttttagcagaagacaaagcaccacattcgaaaga 705
DB 385 aaacgcggcggtatcttctgttcttctgttagaggtcaaatgcagaacacatttgagaa 444

```

```

QY 706 gccgcattcaatttcagtttggaagaagtcagtaacataattgaaccaatagttgttc 765
DB 445 gcaacttacgacactcaaggtggagatataagcagatatgttcgacagacagtgagtc 504
QY 766 catatccctccaaagaacagagataa 789
DB 505 cacatcatgaagaagacagcttaa 528

```

RESULT 5
T68888 standard; DNA; 1014 BP.

AC T68888;

DT 09-DEC-1997 (first entry)

DE NIMA-interacting protein pin1 DNA.

KW pin1; protein interacting with NIMA; cell proliferation; mitosis; peptidyl-prolyl cis/trans isomerase; adenocarcinoma; cancer;

KW leukemia; psoriasis; pemphigus vulgaris; rheumatoid arthritis; acute respiratory distress syndrome; septic shock; inflammation; therapy; ss.

OS Homo sapiens.

FT Key Location/Qualifiers

FT CDS 25..516

FT tag= a

PN W0917986-A1.

PD 22-MAY-1997.

PF 28-OCT-1996; 96WO-US17334.

PR 13-NOV-1995; 95US-0555912.

PA (SALK) SALK INST BIOLOGICAL STUDIES.

PI Hunter T, Lu KP;

PT Protein, pin1, interacting with NIMA - used for treating cell proliferative disorders

PS Claim 8; Page 49-50; 73pp; English.

CC This DNA sequence codes for human pin1 (W18312), an 18 kDa protein that has peptidyl-prolyl cis/trans isomerase activity, associates with NIMA protein kinase, inhibits the mitosis-promoting function of NIMA when overexpressed, and induces mitotic arrest and nuclear fragmentation when depleted. It was identified using a yeast two-hybrid system with *Aspergillus nidulans* nima as bait and cDNA from an HeLa library. A recombinant expression vector comprising the DNA sequence and host cells containing the vector are claimed. Methods are also claimed for identifying proteins that inhibit the mitosis promoting function of NIMA protein kinase and for controlling the growth of a cell by reducing pin1 activity or pin1 expression using an inhibitor, anti-pin1 antibody, antisense nucleotide sequence or ribozyme, or by increasing pin1 activity in the presence of an activator or increasing pin1 expression using an enhancer. This allows treatment of cell proliferation disorders such as adenocarcinomas, cancers, psoriasis, pemphigus vulgaris, acute respiratory distress syndrome, rheumatoid arthritis, septic shock and inflammation.

SO Sequence 1014 BP; 220 A; 310 C; 324 G; 160 T; 0 other;

Query Match 5.4%; Score 53.6; DB 18; Length 1014;
Best Local Similarity 48.4%; Pred. No. 0.00026;
Matches 180; Conservative 0; Mismatches 189; Indels 3; Gaps 1;
QY 418 aacatgttcaagccactgtgaatgagatgagccaggttagatgttcctcattgtg 477
DB 148 agcagtggtgcaaaaacagggcagggggagccctgcagggctcgtcgcactgtcg 207
QY 478 atcaagaacaatcaatcaagaagaacccaagtcttggaagtcgccagatggtataga 537
DB 208 gtgaagcagcagcagtcacggcgccctcgtccgcg---caggagaagatcacccgg 264
QY 538 actagaagcgaatctatacagatattgaagaacatttggaagaataattgagtgtgag 597
DB 265 accaagaaggagccctgagctgcatcaacgcgtacatccagaagaatgctcgagag 324
QY 598 gttaaactgaatgttggaataatccgaagtgtattgagctacatgacagaggtgtg 657
DB 325 gaggaacttgagctctgctcaggtccaggtcagcagctgacgtacgaagcaggcaggga 384
QY 658 gattagaggtttttagcaagaagaacaatgaaccacattcgaagaagccgcatlcaat 717
DB 385 gacctggtgctcttcacagagaggtcagatgacgaagcatttgagaagcctgtttgag 444
QY 718 ttgcattgtgagagaagtcagtaacataatgaaacaaatagtggttcatactccaa 777
DB 445 ctgcgagcggggagatgagcgggcccggtgtcagcagattccgcgcatccatcatcctc 504
QY 778 agaacagagataa 789
DB 505 cgcactgagctga 516
RESULT 6
F07816
ID F07816 standard; cDNA; 662 BP.
AC F07816;
XX
DT 13-MAR-2001 (first entry)
XX
DE Fusarium venenatum EST SEQ ID NO:339.
XX
KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Fusarium venenatum.
XX
PN MO200056762-A2.
XX
PD 28-SEP-2000.
XX
PF 22-MAR-2000; 2000WO-US0781.
XX
PR 22-MAR-1999; 99US-0273623.
XX
PA (NOVO) NOVO NORDISK INC.
XX
PI (NOVO) NOVO NORDISK AS.
XX
PI Berka RM, Rey MM, Shuster JR, Kauppinen S, Clausen IG, Olsen PB;
XX
DR WPI: 2000-594572/56.
XX
PT Monitoring differential expression of genes in filamentous fungal cells
XX uses fluorescence-labeled nucleic acids isolated from the cells and a
XX substrate of expressed sequence tags -
XX
PS Claim 86; Page 514; 3161pp; English.
XX
CC The present invention describes a method for monitoring differential

CC expression of genes in a first filamentous fungal (FF) cell relative to
CC expression of the same genes in one or more second filamentous fungal
CC cells. The method uses fluorescence-labeled nucleic acids isolated from
CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
CC are used in the methods for monitoring differential expression of genes
CC in a first filamentous fungal (FF) cell relative to expression of the
CC same genes in one or more second filamentous fungal cells. Monitoring
CC the global expression of genes from FF cells allows the production
CC potential of the microorganisms to be improved. New genes may be
CC discovered, possible functions of unknown open reading frames can be
CC identified and gene copy number variation and stability can be
CC monitored. The expression of genes can be used to study how FF cells
CC adapt to changes in culture conditions, environmental stress, spore
CC morphogenesis, recombination, metabolic or catabolic pathway engineering.
CC Using ESTs provides several advantages over genomic or random cDNA
CC clones including elimination of redundancy as one spot on an array
CC equals one gene or open reading frame, and organisation of the
CC microarrays based on function of the gene products to facilitate
CC analysis of the results. F07478 to F1247 represents ESTs from Fusarium
CC venenatum; F11248 to F11853 represents ESTs from Aspergillus niger;
CC F11854 to F14878 represents ESTs from Aspergillus oryzae; and F14879 to
CC F15337 represents ESTs from Trichoderma reesei, which are all
CC specifically claimed in the present invention.
XX
SQ Sequence 662 BP; 169 A; 190 C; 159 G; 141 T; 3 other:
Query Match 5.4%; Score 53.2; DB 21; Length 662;
Best Local Similarity 50.0%; Pred. No. 0.00029;
Matches 248; Conservative 0; Mismatches 233; Indels 15; Gaps 4;
QY 271 acaggtaccacctaattgacgattagatgacatccagtcacataaagaatattc 330
DB 119 accggtcttcctgctgtggtgaggtgcgaactccaactccaagaacccctccattac 178
QY 331 ttaaccaatctaccatgagtgctcttgaccacctatgagcactgcaagaagta 390
DB 179 ttcaactcgcgaagaagctgtcgagatggaaacccctccgcagcagacaggaag 238
QY 391 ttgaatgatacattgcgaagttaaataaacaatgttacaaagcacttgyaatgagat 450
DB 239 ctcaagcactacatgycgaacgaacataagcgtgttcgcgtcctgtgtgcagttcccg 298
QY 451 ggcacag-----ttagagttctcaattgttgatcaagaacaatcaatcaaga 501
DB 299 gtcccagagggcaaatctgcgcacatctctcgtcaagcatcgcgacgcagcg 358
QY 502 cccaagcttggaagtcaccaagatggtataagtgtaactgagcagatcatacagata 561
DB 359 cccaacagctlgagagagcgca--gattactgcgtccaagagagaggtcttcagatc 415
QY 562 ttgaagaacaattggaagaatattgagtgtgaggttaactaaagtgaat-tggcaaa 620
DB 416 atcaagagagcagcaagaagaatcaagtcgtgaagtgtcagccttgcgaagctgtctc 475
QY 621 taccgaagtgatgcagctacacatgacagagtggtgattaggtttttagcaa--a 678
DB 476 caccgagtcgcagctgcctcactcgaagcgttggtggatttggaacttttgcaagg 535
QY 679 ggacaatgcaaccacattcgaagaagccgcatlcaatttgcatgttggaagaatcagt 738
DB 536 ggagatagcaaaaggttggaagatgcttcttcttcttcaagccacatgagc 595
QY 739 aacataattgaacca 754
DB 596 gagatgttgaaacaa 611
RESULT 7
F11336
ID F11336 standard; cDNA; 406 BP.
XX
AC F11336;

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 9547.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139482.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.9%; Score 38.6; DB 21; Length 666;
Best Local Similarity 52.1%; Pred. No. 1.4;
Matches 86; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 630 tgaatgagctacatgacagaggtgtgattgaagggttttttagcaaggacaatgca 689
DB 343 ttaggtgtccatcagggaaaaagagagatcttgaatgtgtccctcgtgaggaagatgac 402
QY 690 accacattgagaagacgcatcatttgcattgttggaggaagtgcagtaacataatga 749
DB 403 aggtccattccaaagatgtcgcccttaatacacactgttggaggtcacaggtgcacccctcaa 462
QY 750 aaccaatagtgtgtgcataatcctccaagaagacaagataatcaa 794
DB 463 atcaagcagcagatcacacacatttattatcagagggagaagaaga 507

RESULT 11
X20248 standard; DNA; 910715 BP.
AC X20248;
XX 04-MAY-1999 (first entry)
DT

XX XX
DE Borrelia burgdorferi polynucleotide sequence #1.
XX
KW Borrelia burgdorferi; spirochete; bacterium; pathogen; Lyme disease;
KW epidemic relapsing fever; endemic relapsing fever; Lyme borreliosis;
KW infection; diagnosis; characterisation; detection; ds.
XX
OS Borrelia burgdorferi.
XX
PN W09858943-A1.
XX
PD 30-DEC-1998.
XX
PF 18-JUN-1998; 98WO-US12764.
XX
PR 03-SEP-1997; 97US-0057483.
PR 20-JUN-1997; 97US-0050359.
PR 22-JUL-1997; 97US-0053344.
PR 22-JUL-1997; 97US-0053377.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (MED-) MEDIMUNE INC.
PI Clayton R, Dougherty BA, Fraser C, Lathigra R, Smith HO;
PI White OR;
XX WPI; 1999-081217/07.
XX
XX New isolated Borrelia burgdorferi nucleic acids - used to develop
PT products for the detection, diagnosis, characterisation, prevention
PT and therapy of infections, particularly Lyme disease
XX
PS Claim 1; Page 157-671; 1128pp; English.
XX
CC X20248 to X20402 represent polynucleotide sequences isolated from
CC Borrelia burgdorferi (Bb). Products derived from Bb can be used for
CC the detection, diagnosis, characterisation, prevention and therapy of
CC Bb infections, e.g. Lyme disease. They can also be used for the
CC production of biosynthetic products, e.g. enzymes. Borrelia belongs
CC to a family of motile, spiral-shaped bacteria called Spirochetes.
CC Spirochetes are pathogenic in humans and Borrelia causes epidemic and
CC endemic relapsing fever, and Lyme borreliosis, more commonly known as
CC Lyme disease.
XX
SQ Sequence 910715 BP; 327171 A; 129646 C; 130753 G; 323091 T; 54 other;

Query Match 3.9%; Score 38.4; DB 20; Length 910715;
Best Local Similarity 52.5%; Pred. No. 15;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 701 aagaagccgcatcatttgcattgttgagagatgcagtaacataatgaaacaaatgtg 760
DB 383284 aaataagactcaactaaatgtctctgtatctcagagcatcatttaatttttgaatcaaacgac 383343
QY 761 gtgtccatctccccaagaacagataaatacaagatattggaattgtgataaatagaa 820
DB 383344 aatlaaatctacttaagaatlaaatatataataaactcatttaaaaaaacaacaaca 383403
QY 821 aataataagagacaagttgtatagattgtgtacccaanaa 860
DB 383404 tcttaataaaaactattttaaaaatttcctaactaca 383443

RESULT 12
C62126
ID C62126 standard; DNA; 4468 BP.
XX
AC C62126;
XX 06-MAR-2001 (first entry)
DT
XX DNA encoding a calcium-dependent serine-protease designated Pf-SUB2.
DE

XX Calcium-dependent serine-protease; Pf-SUB2; merozoite differentiation;
 KW major surface protein 1; MSP1-42; erythrocyte entry; malaria; ss.
 XX Plasmodium falciparum.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 61..4074
 FT /*tag= a
 FT /product= "serine protease"
 XX
 PN FR2791685-A1.
 PD 06-OCT-2000.
 XX
 PF 31-MAR-1999; 99FR-0004039.
 XX
 PR 31-MAR-1999; 99FR-0004039.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Barale JC, Langsley G, Braun BC, Pereira Da Silva L, Blisnick T;
 DR WPI; 2000-658021/64.
 DR P-PSDB; B30504.
 XX
 PT Polypeptide with calcium-dependent serine-protease activity, for the
 PT prevention, treatment, and detection of malarial infections due to
 PT Plasmodium falciparum
 XX
 PS Claim 3; Page 33-34; 47pp; French.
 XX
 CC The present sequence encodes a polypeptide (Pf-SUB2) which has a
 CC calcium-dependent serine-protease activity. The Pf-SUB2 gene is
 CC expressed during the differentiation phase of merozoites. The protein
 CC is implicated in maturation of the major surface protein 1 of
 CC merozoites (MSP1-42). The enzyme is also crucial for entry of the
 CC parasite into erythrocytes. The polypeptides and polynucleotides are
 CC used to identify inhibitors of Pf-SUB2. These inhibitors e.g. antibodies,
 CC are used for the detection, prevention, and treatment of malaria due to
 CC Plasmodium falciparum infection.
 CC
 XX
 SQ Sequence 4468 BP; 1902 A; 456 C; 641 G; 1469 T; 0 other;
 Query Match 3.9%; Score 38.2; DB 21; Length 4468;
 Best Local Similarity 52.1%; Pred. No. 3.2;
 Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 XX
 OY 74 attagcaactggaatagatatagagttgagctgcttgcgaatagatattagtgatg 133
 Db 4248 attatataatataatataatataatataatataatataatataatattagatat 4307
 OY 134 tacattacaacaaactctctcttcttcataatcttcaacaacaagaatttcgtgt 193
 Db 4308 ttatttttaaatagtaactcatttttaagtaaacacattccctcttttcgttt 4367
 OY 194 tgcctttgtgtattatttgcacatcgcttaagcttgcattc 236
 Db 4368 ttcatattattattattattattattattatttttttttt 4410
 XX
 RESULT 13
 ID C62127 standard; DNA; 4611 BP.
 XX
 AC C62127;
 XX
 DT 06-MAR-2001 (first entry)
 XX
 DE DNA encoding a calcium-dependent serine-protease designated Pf-SUB2.
 XX

KW Calcium-dependent serine-protease; Pf-SUB2; merozoite differentiation;
 KW major surface protein 1; MSP1-42; erythrocyte entry; malaria; ss.
 XX Plasmodium falciparum.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 124..4217
 FT /*tag= a
 FT /product= "serine protease"
 FT /note= "contains an intron"
 FT exon 124..2877
 FT /*tag= b
 FT Intron 2878..2930
 FT /*tag= c
 FT exon 2931..4217
 FT /*tag= d
 XX
 PN FR2791685-A1.
 PD 06-OCT-2000.
 XX
 PF 31-MAR-1999; 99FR-0004039.
 XX
 PR 31-MAR-1999; 99FR-0004039.
 XX
 PA (INSP) INST PASTEUR.
 PA (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Barale JC, Langsley G, Braun BC, Pereira Da Silva L, Blisnick T;
 DR WPI; 2000-658021/64.
 DR P-PSDB; B30505.
 XX
 PT Polypeptide with calcium-dependent serine-protease activity, for the
 PT prevention, treatment, and detection of malarial infections due to
 PT Plasmodium falciparum
 XX
 PS Claim 3; Page 34-45; 47pp; French.
 XX
 CC The present sequence encodes a polypeptide (Pf-SUB2) which has a
 CC calcium-dependent serine-protease activity. The Pf-SUB2 gene is
 CC expressed during the differentiation phase of merozoites. The protein
 CC is implicated in maturation of the major surface protein 1 of
 CC merozoites (MSP1-42). The enzyme is also crucial for entry of the
 CC parasite into erythrocytes. The polypeptides and polynucleotides are
 CC used to identify inhibitors of Pf-SUB2. These inhibitors e.g. antibodies,
 CC are used for the detection, prevention, and treatment of malaria due to
 CC Plasmodium falciparum infection.
 CC
 XX
 SQ Sequence 4611 BP; 1951 A; 459 C; 657 G; 1544 T; 0 other;
 Query Match 3.9%; Score 38.2; DB 21; Length 4611;
 Best Local Similarity 52.1%; Pred. No. 3.2;
 Matches 85; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
 XX
 OY 74 attagcaactggaatagatatagagttgagctgcttgcgaatagatattagtgatg 133
 Db 4391 attatataatataatataatataatataatataatataatataatattagatat 4450
 OY 134 tacattacaacaaactctcttcttcataatcttcaacaacaagaatttcgtgt 193
 Db 4451 ttatttttaaatagtaactcatttttaagtaaacacattccctcttttcgttt 4510
 OY 194 tgcctttgtgtattatttgcacatcgcttaagcttgcattc 236
 Db 4511 ttcatattattattattattattattattatttttttttt 4553
 XX
 RESULT 14
 ID Z01425 standard; DNA; 1038602 BP.
 XX

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2001, 03:58:30 ; Search time 1747.82 Seconds
(without alignments)
4943.301 Million cell updates/sec

Title: us-09-507-242-1

Perfect score: 989

Sequence: 1 gatcaacacatagatgtgt.....ggtttgtcatattgtgatcc 989

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 seqs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

10: gb_est10:*

11: gb_est11:*

12: gb_est12:*

13: gb_est13:*

14: gb_est14:*

15: gb_est15:*

16: gb_est16:*

17: gb_est17:*

18: gb_est18:*

19: gb_est19:*

20: gb_est20:*

21: gb_est21:*

22: gb_est22:*

23: gb_est23:*

24: gb_est24:*

25: gb_est25:*

26: gb_est26:*

27: gb_est27:*

28: gb_est28:*

29: gb_est29:*

30: gb_est30:*

31: gb_est31:*

32: gb_est32:*

33: gb_est33:*

34: gb_est34:*

35: gb_est35:*

36: gb_est36:*

37: gb_est37:*

38: gb_est38:*

39: gb_est39:*

40: gb_est40:*

41: gb_est41:*

42: gb_est42:*

43: gb_est43:*

44: em_esthum10:*

45: em_esthum11:*

46: em_esthum12:*

47: em_esthum13:*

48: em_esthum14:*

49: em_esthum15:*

50: em_esthum16:*

51: em_esthum17:*

52: em_esthum18:*

53: em_esthum19:*

54: em_esthum20:*

55: em_esthum21:*

56: em_esthum22:*

57: em_esthum23:*

58: em_esthum24:*

59: em_esthum25:*

60: em_esthum26:*

61: em_esthum27:*

62: em_esthum28:*

63: em_esthum29:*

64: em_esthum30:*

65: em_esthum31:*

66: em_esthum32:*

67: em_esthum33:*

68: em_esthum34:*

69: em_esthum35:*

70: em_esthum36:*

71: em_esthum37:*

72: em_esthum38:*

73: em_esthum39:*

74: em_esthum40:*

75: em_esthum41:*

76: em_esthum42:*

77: em_esthum43:*

78: em_esthum44:*

79: em_esthum45:*

80: em_esthum46:*

81: em_esthum47:*

82: em_esthum48:*

83: em_esthum49:*

84: em_esthum50:*

85: em_esthum51:*

86: em_esthum52:*

87: em_esthum53:*

88: em_esthum54:*

89: em_esthum55:*

90: em_esthum56:*

91: em_esthum57:*

92: em_esthum58:*

93: em_esthum59:*

94: em_esthum60:*

95: em_esthum61:*

96: em_esthum62:*

97: em_esthum63:*

98: em_esthum64:*

99: em_esthum65:*

100: em_esthum66:*

101: em_esthum67:*

102: em_esthum68:*

103: em_esthum69:*

104: em_esthum70:*

105: em_esthum71:*

106: em_esthum72:*

107: em_esthum73:*

108: em_esthum74:*

109: em_esthum75:*

110: em_esthum76:*

111: em_esthum77:*

112: em_esthum78:*

113: em_esthum79:*

114: em_esthum80:*

115: em_esthum81:*

116: em_esthum82:*

117: gb_est48:*
 118: gb_est49:*
 119: gb_est50:*
 120: gb_est51:*
 121: gb_est52:*
 122: gb_est53:*
 123: gb_est54:*
 124: gb_est55:*
 125: gb_est56:*
 126: gb_est57:*
 127: gb_est58:*
 128: gb_est59:*
 129: gb_est60:*
 130: gb_est61:*
 131: gb_est62:*
 132: gb_est63:*
 133: gb_est64:*
 134: gb_est65:*
 135: gb_est66:*
 136: gb_est67:*
 137: gb_est68:*
 138: gb_est69:*
 139: gb_est70:*
 140: gb_est71:*
 141: gb_est72:*
 142: gb_est73:*
 143: gb_est74:*
 144: gb_est75:*
 145: gb_est76:*
 146: gb_est77:*
 147: gb_est78:*
 148: gb_est79:*
 149: gb_est80:*
 150: gb_est81:*
 151: gb_est82:*
 152: gb_est83:*
 153: gb_est84:*
 154: gb_est85:*
 155: gb_est86:*
 156: gb_est87:*
 157: gb_est88:*
 158: gb_est89:*
 159: gb_est90:*
 160: gb_est91:*
 161: gb_est92:*
 162: gb_est93:*
 163: gb_est94:*
 164: gb_est95:*
 165: gb_est96:*
 166: gb_est97:*
 167: gb_est98:*
 168: gb_est99:*
 169: gb_est100:*
 170: gb_est101:*
 171: gb_est102:*
 172: gb_est103:*
 173: gb_est104:*
 174: gb_est105:*
 175: gb_est106:*
 176: gb_est107:*
 177: gb_est108:*
 178: gb_est109:*
 179: gb_est110:*
 180: gb_est111:*
 181: gb_est112:*
 182: gb_est113:*
 183: gb_est114:*
 184: gb_est115:*
 185: gb_est116:*
 186: gb_est117:*
 187: gb_est118:*
 188: gb_est119:*
 189: gb_est120:*

190: em_gss_pln1:*
 191: em_gss_pln2:*
 192: em_gss_pro:*
 193: em_gss_rod1:*
 194: em_gss_rod2:*
 195: em_gss_rod3:*
 196: em_gss_rod4:*
 197: em_gss_rod5:*
 198: em_gss_vrt1:*
 199: em_gss_vrt2:*
 200: em_gss_vrt3:*
 201: gb_gss1:*
 202: gb_gss2:*
 203: gb_gss3:*
 204: gb_gss4:*
 205: gb_gss5:*
 206: gb_gss6:*
 207: gb_gss7:*
 208: gb_gss8:*
 209: gb_gss9:*
 210: gb_gss10:*
 211: gb_gss11:*
 212: gb_gss12:*
 213: gb_gss13:*
 214: gb_gss14:*
 215: gb_gss15:*
 216: gb_gss16:*
 217: gb_gss17:*
 218: gb_gss18:*
 219: gb_gss19:*
 220: gb_gss20:*
 221: gb_gss21:*
 222: gb_gss22:*
 223: gb_gss23:*
 224: gb_gss24:*
 225: gb_gss25:*
 226: gb_gss26:*
 227: gb_gss27:*
 228: gb_gss28:*
 229: gb_gss29:*
 230: gb_gss30:*
 231: gb_gss31:*
 232: gb_gss32:*
 233: gb_gss33:*
 234: gb_gss34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	101.2	10.2	569	121	AM669911	AM669911 NXNV_123-
2	100.2	10.1	480	138	BE662291	BE662291 STR4/STR4
3	99.4	10.1	601	110	AM064606	AM064606 STR3/STR10 P
4	96.4	9.7	412	32	AM010930	AM010930 STR13/12 P
5	96	9.7	615	114	AM334625	AM334625 S3BD12 AG
6	94.8	9.6	462	137	BE607245	BE607245 NXCT_034-
7	94.2	9.5	471	106	AU051808	AU051808 AU051808
8	94.2	9.5	517	175	C92940	C92940 C92940 Dict
9	94.2	9.5	659	175	C92104	C92104 C92104 Dict
10	92	9.3	729	106	AU000273	AU000273 AU000273
11	90.6	9.2	708	32	AM010527	AM010527 STR7/611 P
12	90.4	9.1	738	29	AV400559	AV400559 AV400559
13	90.4	9.1	748	106	AU003598	AU003598 AU003598
14	89.8	9.1	660	106	AU000274	AU000274 AU000274
15	89.8	9.1	722	106	AU003857	AU003857 AU003857
16	89.8	9.1	731	106	AU003742	AU003742 AU003742
17	89.8	9.1	740	106	AU003792	AU003792 AU003792
18	89.6	9.1	630	175	C90182	C90182 C90182 Dict

```

19 87.2 8.8 703 162 BE038350
20 81.8 8.3 577 24 A1772554
21 81.8 8.3 577 173 BG133271
22 81.8 8.3 648 122 AM945046
23 80.6 8.1 699 102 A1822163
24 80.6 8.1 780 161 BE034276
25 76.4 7.7 492 122 BE034276
26 76.4 7.7 492 122 BE034276
27 76.4 7.7 587 142 BE003579
28 76.4 7.7 671 114 AM397670
29 76.4 7.7 671 114 AM397670
30 73.2 7.4 665 110 AM054423
31 73.2 7.4 437 21 A1496600
32 72.7 7.3 356 127 T76444
33 72.7 7.3 810 146 BE280096
34 71.8 7.3 467 165 BE346524
35 71.8 7.3 511 120 AM831738
36 71.4 7.2 497 165 BE316431
37 70.8 7.2 513 105 AL387506
38 70.8 7.2 205 111 AM145706
39 70.4 7.1 960 162 BE054767
40 70.2 7.1 452 163 BE190190
41 70.2 7.1 503 146 BE324477
42 70.2 7.1 261 141 BE921583
43 69.8 7.1 702 24 A1730273
44 68 6.9 400 119 AM761425
45 67.8 6.9 642 162 BE037023

```

ALIGNMENTS

```

RESULT 1
LOCUS AM869911 569 bp mRNA EST 22-MAY-2000
DEFINITION NXNV_123_A01.F Nsf Xylem Normal wood Vertical Pinus taeda cDNA
VERSION AM869911
KEYWORDS AM869911.1 GI:8004060
SOURCE EST.
ORGANISM loblolly pine.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 569)
AUTHORS Sederoff, R.
TITLE Molecular Basis of Wood Formation in the Pine Megagenome
JOURNAL Unpublished (2000)
COMMENT Contact: Johnson, Arthur
North Carolina State University
Tel: 919 515 7800
Fax: 919 515 7801
Email: ajohnson@unity.ncsu.edu
Seq primer: T3.
FEATURES
source
1..569
Location/Qualifiers
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="NXNV_123_A01"
/note="Vector: Bluescript SK; Site_1: Eco RI; The
sequences contain a 'CDNA adapter' between the EcoRI site
and the start of the EST. The adapter sequence is
'AATTGGCAGCAGC'."
BASE COUNT 172 a 105 c 131 g 140 t 21 others
ORIGIN

```

Query Match 10.2%; Score 101.2; DB 121; Length 569;
 Best Local Similarity 58.1%; Pred No. 2.4e-14;
 Matches 200; Conservative 0; Mismatches 135; Indels 9; Gaps 1;

Qy 455 aggttagagttctcatctgtgtatcaagaacatcaatcaagaacaaacccaagttcttga 514

```

Db 127 AGGTGAGAGCTTCTCACTTGCTTATCAACAGAGGATCAAGAGGAGCTTCTCTGTC
515 agtccccaataggtataag-----tagaactagaacgaatctatcagatatga 565
187 AGACCCAGATGCGCCAGAGATTAAAGCTAACCCAGAGATGCTGCGTGGCCCACTCA 246
566 agaaacatttgaaagaacatcttgatggtgaggttaactaagtgaattggcaaatccg 625
247 GTCATTTGAGGAGAGAAATTTGTTAGTGAAGACCAAGTTTAGACCTTGGCGTCCGAT 306
626 aaagttagtcagctcacatgacagagtggtgattaggtttttagcaaaagacaa 685
307 ATTCCGACTGTAAATCCGCTTAAGAAAGCGGAGATTGGTCCATTTGGTCGAGTCAAA 366
686 tgcgaaccaccattcgaaagacgcgcattcaattgcatgttgaggaagtcagtaataa 745
367 TGCAGAAAGGCTTTCGAAAGTGTCTACATACCTTTTGAAGTGGCGAGATTGATGATTTG 426
746 tgaacacaaatagtggtgctcatatccccaagaacagatata 789
427 TGCACACCGACAGTGTGTGTCATATCATCTTAANNACAGATCA 470

```

```

RESULT 2
LOCUS BE662291 480 bp mRNA EST 08-SEP-2000
DEFINITION ST84/ST84C02 Pine Triplex shoot tip library Pinus taeda cDNA clone
VERSION BE662291
KEYWORDS BE662291.1 GI:9993028
SOURCE EST.
ORGANISM loblolly pine.
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 480)
AUTHORS Whetten, R.W., Kinlaw, C.S., Retzel, E. and Sederoff, R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhetten@unity.ncsu.edu
Seq primer: 5' lambda triplex2 Sequencing Primer.
FEATURES
source
1..480
Location/Qualifiers
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="ST84/ST84C02"
/clone_lib="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1:
SfiI (A); Site_2: SfiI (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT 138 a 85 c 119 g 122 t 16 others
ORIGIN

```

Query Match 10.1%; Score 100.2; DB 138; Length 480;
 Best Local Similarity 57.8%; Pred No. 4.1e-14;
 Matches 199; Conservative 0; Mismatches 136; Indels 9; Gaps 1;

Qy 455 aggttagagttctcatctgtgtatcaagaacatcaatcaagaacaaacccaagttcttga 514

```

Db 38 AGGTGAGAGCTTCTACCTTGGCTTATCAAAACGAGGATCAAGAGGCGCTTCCTCGTGGC 97
Qy 515 agtcccaagatggtataga-----tagaactagaagaactctatacagatattga 565
Db 98 AAGACCCAGATGCGCGAAGGATTAAAGCTAACACCAAGATGCTGCCGTTCCCAACTCA 157
Qy 566 agaaacattggaagaatatagtggtgaggttaaacactaagtaagtaattggcaataccg 625
Db 158 GTGCATTGAGGAGGAATAATTGTAGTGAAGGCCAAGTTGAGGACCTTCGGCTCGAT 217
Qy 626 aaagtattgacgtccacatgacagaggtggtgatttagggtttttagcaagaagacaa 685
Db 218 ATTCGCACTGTAATTCGCTTAAGAAAGGNGNATTTGGGTCATTGGTCGACGTCAAA 277
Qy 686 tgcacacacattcgaaagccgacattcaattgcatggttggaggaagtcagtaacataa 745
Db 278 TGCAAAAGCCCTTTGAAGATCTCTACATACCTTTTGAAGTTGGCGAGATTGATGATTTG 337
Qy 746 ttgaaccaatagtggtgtccatactctccaagaagacagataa 789
Db 338 TGGACACTGACAGTGTGTGCATATCATCTTAAGACGACGATGA 381

```

```

RESULT 3
AM064606 601 bp mRNA EST 12-OCT-1999
LOCUS ST33E10, Pinus Triplex shoot tip library Pinus taeda cDNA clone
DEFINITION ST33E10, mRNA sequence.
ACCESSION AM064606
VERSION AM064606.1 GI:6019678
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

```

```

REFERENCE
AUTHORS Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
,NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.

```

```

FEATURES
source
1.601
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="SP33E10"
/clone_1bp="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1:
Site_1 (A); Site_2: Site_1 (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT 189 a 113 c 147 g 150 t 2 others
ORIGIN

```

```

Query Match 10.1%; Score 99.4; DB 110; Length 601;
Best Local Similarity 56.9%; Pred. No. 6.6e-14;
Matches 206; Conservative 0; Mismatches 147; Indels 9; Gaps 1;
Qy 455 aggttagagttctcattgttgatcaagaacaatcaatcaagaagaacgaagtccttga 514

```

```

Db 160 AGGTGAGAGCTTCTACCTTGGCTTATCAAAACGAGGATCAAGAGGCGCTTCCTCGTGGC 219
Qy 515 agtcccaagatggtataga-----tagaactagaagaactctatacagatattga 565
Db 220 AAGACCCAGATGCGCGAAGGATTAAAGCTAACACCAAGATGCTGCCGTTCCCAACTCA 279
Qy 566 agaaacattggaagaatatagtggtgaggttaaacactaagtaagtaattggcaataccg 625
Db 280 GTGCATTGAGGAGGAATAATTGTAGTGAAGGCCAAGTTGAGGACCTTCGGCTCGAT 339
Qy 626 aaagtattgacgtccacatgacagaggtggtgatttagggtttttagcaagaagacaa 685
Db 340 ATTCGCACTGTAATTCGCTTAAGAAAGGCGGAGATTGGGTCATTGGTCGACGTCAAA 399
Qy 686 tgcacacacattcgaaagccgacattcaattgcatggttggaggaagtcagtaacataa 745
Db 400 TGCAAAAGCCCTTTGAAGATCTCTACATACCTTTTGAAGTTGGCGAGATTGATGATTTG 459
Qy 746 ttgaaccaatagtggtgtccatactctccaagaagacagataaagatattgagt 805
Db 460 TGGACACTGACAGTGTGTGCATATCATCTTAAGACGACGATGA 519
Qy 806 tt 807
Db 520 TT 521

```

```

RESULT 4
AM010930 412 bp mRNA EST 10-SEP-1999
LOCUS ST13C12, Pinus Triplex shoot tip library Pinus taeda cDNA clone
DEFINITION ST13C12, mRNA sequence.
ACCESSION AM010930
VERSION AM010930.1 GI:5859624
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

```

```

REFERENCE
AUTHORS Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
,NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unity.ncsu.edu
Seq primer: 5' lambda Triplex2 Sequencing Primer.

```

```

FEATURES
source
1.412
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="SP13C12"
/clone_1bp="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site_1:
Site_1 (A); Site_2: Site_1 (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT 115 a 82 c 107 g 104 t 4 others
ORIGIN

```

```

Query Match 9.7%; Score 96.4; DB 32; Length 412;

```

[illegible][illegible]

537 CATTCTTCCGAC--CGACCTATTATCCGACGACCTATGATATTATTC 481

566 aaaaaacattttgaaaaaataattaaatontaaattaaaactaaataaatgaaattgacaaataccg 625

Db 480 TTAAATATGAGGC AAAAGTTGCTCAGGAGACATTAGTCTGGGAACTAGCTATTACGG 421

626 aaataattacaactcacatgacagaagtqatatttaaatTTTtaacaaagacaaa 685

Db 420 AATCGATTGTAATAGTCTAAAAATCGGGGATTAGCTTTTGGAAAGGTGTTA 361

686 tqcaaccacccattcgaagaagccgcattcaatttgcattgttggaagaagtcagtaacataa 745

Db 360 TGCACGGGAATTGAGGAGGCATCGTTAGTTTACAACGAGAGATGAGTCATGTAG 302

QY 746 ttgaaccaatagtggtgtcattatcctccaagaacagataataatcaagatatgtgagt 805

Db 300 TAGAACTGCTAGTGGTTATCATTTAATTGAAGAATTGCTTAAGATCTTTGAATGATT 241

QY 806 ttgatgaa 813

Db 240 TTATTTAA 233

РЕСНИК 5

Accession	Gene	Length (bp)	Source	EST	Accession
BE607245	LOC115483	462	mRNA	EST	22-ATIG-2000

DEFINITION	DESCRIPTION
NXCI_034_E09_F	NXCI (Nsi Xylem compression wood Inclined) Pinus taeda cDNA clone NXCI 034 E09 5' mRNA sequence.

ACCESSION	BE607243
VERSION	BE607245.1
	GI:9884325

NEWWORDS	Est.
SOURCE	loblolly pine.

OKONNASHI
FINUS CAUDA
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta

REFERENCE 1 (bases 1 to 462)

TITLE
Molecular Basis of Wood Formation in the Pine Megagenome

COMMENT
Contact: Johnson, Arthur

Tel: 919 515 7800

Email: ajohnson@unity.ncsu.edu

FEATURES	Location/Qualifiers
1	1453

```
/organism="Pinus taeda"
/strain="Goccal plant 10b101" from North Carol
```

```

/db_xref="taxon:3352"
/clone="NYCT 034 F00"

```

```
/clone_lib="NXCI (Nsf xylem compression wood Inclined  
/tissue_type="Xylem"
```

```
/cell_type="compression"
/dev stage="Juvenile"
```

```
/lab_host="XLI-Blue"
/note="Vector: Bluescript SK: Site 1: Eco RI: Site 2:
```

; the library is from early (spring) wood, taken from three six-year old trees (three different genotypes),

compression wood by bending to a 45 degree angle and the juvenile phase. These crees were induced to form

them to the ground. Differentiating system was induced from the bottoms of the inclined stems, and a mixture

primed cDNA was directionally cloned into the EcoRI-X

'cdna adapter' between the *EcoRI* site and the start o

QY	588	gagtggtgaggtgtaaaccaaggaagtgatctggcaatcccgaaagtgatctggagctcacatga	647
Db	206	ATCTGCTTCGCGACATTTGGAAGATTGACCTATATAAATTCAGATTGATGACAGCTCTAA	265
QY	648	cagaagtggtgatttagggcttttttagcaaaaggaacaaatgcaaccacatctcgaaagac	707
Db	266	AAGAGTGCGTATCTTGTATCCATTCATAAAGAGCCCAAAATGCAGACCTTTGGAAGATTG	325
QY	708	cgcatacattgcatctgttgagaagatcagtcacataatctgaaccaatagtgtgtcca	767
Db	326	CGCATTTCTTTTAAAGCTTGGAAGATTAGTGCGCATTTGTATGATAGTATGATGAGTGCTTCA	385
QY	768	tatctccaaagaacagataatcaatga	798
Db	386	TATTTATTGAAGATTAGCATTAATATAGTATA	416
RESULT	9		
LOCUS	C92104	659 bp	mRNA EST 12-JUL-1999
DEFINITION	C92104 Dictyostelium discoidium SS (H.Urushihara) Dictyostelium		
ACCESSION	C92104		
VERSION	C92104.1	GI:3073980	
KEYWORDS	EST.		
SOURCE	Dictyostelium discoidium.		
ORGANISM	Dictyostelium discoidium		
REFERENCE	Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.		
AUTHORS	1 (bases 1 to 659) Mortio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M., Yoshino,R., Mitra,B.N., Pl.M., Sato,T., Takemoto,K., Yasukawa,H., Williams,J., Maeda,M., Takeuchi,T., Ochiai,H. and Tanaka,Y. The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development DNA Res. 5 (6), 335-340 (1998)		
TITLE			
JOURNAL			
MEDLINE	99156227		
COMMENT	Contact: Hideko Urushihara Institute of Biological Sciences University of Tsukuba 3-3-10 Ten-nodal, Tsukuba, Ibaraki 305, Japan Email: d402huesakura.cc.tsukuba.ac.jp PROJECT = 'Dictyostelium discoidium cDNA project in Japan' POLYA-No.		
FEATURES			
source	Location/Qualifiers 1..659 /organism="Dictyostelium discoidium" /strain="AX4" /db_xref="taxon:44689" /clone="SSD494" /clone_id="Dictyostelium discoidium SS (H.Urushihara)" /dev_stage="slug"		
BASE COUNT	259 a 102 c 88 g 210 t		
ORIGIN			
Query Match	9.5%; Score 94.2; DB 175; Length 659;		
Best Local Similarity	57.4%; Pred. No. 1.2e-12;		
Matches 190; Conservative 0; Mismatches 138; Indels 3; Gaps 1			
QY	468	tcaattgtgatcaagaacaatcaatcaagaanaaccacaagctcttggaagtcgccagatg	527
Db	228	TCATCTTTTGTAGTTAAACATCAAGGTTCAAGAAATCCATCATCTTTGGAGG---GAATCTAA	284
QY	528	tataagtagaactagagacgaatctatacagatcttgagaagaacatttgaaagaatct	587
Db	285	AATTTCAGAAACCAAGAAAGAGCAATGCTTAATCTTAATGATATAGCAACAACTTAT	344
QY	588	gagtggtgaggtttaaactaagtgaaatctggcaaatccgaaagtgatctgagctcacatga	647
Db	345	ATCTGCTTCGCGACATTTGGAAGATTAGCTCATTAATAATTCAGATTGTAGCGAGCTCAA	404
QY	648	cagaagtggtgatttagggcttttttagcaaaaggaacaaatgcaaccacatctcgaaagac	707

[illegible]

OY 882 aaactgt 889
Db 667 AGATCTTT 674

RESULT 11

AM010527 708 bp mRNA EST 10-SEP-1999
LOCUS S707G11 Pine Triplex shoot tip library Pinus taeda cDNA clone
DEFINITION S707G11, mRNA sequence.
ACCESSION AM010527
VERSION AM010527.1 GI:5859305
KEYWORDS EST.
SOURCE loblolly pine.
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.

REFERENCE 1 (bases 1 to 708)
AUTHORS Whetten,R.W., Kinlaw,C.S., Retzel,E. and Sederoff,R.R.
TITLE The Pine Gene Discovery Project
JOURNAL Unpublished (1999)
COMMENT Contact: Ross Whetten
Forest Biotechnology Group
North Carolina State University
Dept. of Forestry, NC State University, 6113 Jordan Hall, Raleigh
NC, 27695-8008
Tel: 919-515-7800
Fax: 919-515-7801
Email: rosswhet@unity.ncsu.edu

Seq primer: 5' Lambda Triplex2 Sequencing Primer.
Location/Qualifiers

FEATURES

source
1. 708
/organism="Pinus taeda"
/db_xref="taxon:3352"
/clone="SF07G11"
/clone_id="Pine Triplex shoot tip library"
/lab_host="E. coli BM25.8"
/note="Organ: shoot tips; Vector: Lambda Triplex; Site: 1;
SfiI (A); Site 2: SfiI (B); Shoot tips (approx. 2 cm from
apex) were collected during the spring, frozen and used
for mRNA isolation. The SMART-PCR method (Clontech) was
used to prepare a library from 1 ug total RNA, using the
Lambda Triplex vector. Plasmid subclones in pTriplex were
recovered by cre-lox excision in E. coli strain BM25.8 and
sequenced from the 5' end."
BASE COUNT 215 a 137 c 165 g 163 t 28 others
ORIGIN

Query Match 9.24; Score 90.6; DB 32; Length 708;
Best Local Similarity 58.34; Pred. No. 9.1e-12;
Matches 201; Conservative 0; Mismatches 134; Indels 10; Gaps 2;

OY 455 aggttagagttcattgttgcataagaacaatcaatcaagaacccaagtcttga 514
Db 128 AGGTCAAGAGCTTCTCATTGCTTATCAACACGAGGATCAAGAGGCGCTTCGTGGC 187
OY 515 agtcccaagatgata-----tagaactagagacgaatctatacagatatga 565
Db 188 AAGACCCAGATGGCCGAAGGATTAAAGCTACACCAAGAGATGTCGCGTGGCCCACTCA 247
OY 566 agaacaatttgaagaatatatgtgtgtgtgaactaagtaagtgaattgcaaataccg 625
Db 248 GTGATGATGAGGAAGAATTTTATGTGGAAGGCCCAAGTTTGAGAGACTTGCAGCTGAT 307
OY 626 aaagttagtgcagctacatgacagagtggtgatttagagtttttagcaaaagacaaa 685
Db 308 ATTCGCACTGTAATTCGCTAAAGAAAGCGAGATTGGGTCATTTGGTGGTGAAGTCAAA 367
OY 686 tgcacaacacattgagaagaagccgacatcaattgcatgttgagaga-agtcaagtaacata 744
Db 368 TGCAAAAGCCTTTTGAAGATGCTACATACCTTTTGAAGATTGGCGAGATTAAAGTATATT 427

OY 745 atgaaaccaatagtggtgtccatatactccaaagaacagatata 789
Db 428 GTGACACAGTACAGTGGTGTGATCATCATCTTAAGACAGATGA 472

RESULT 12

AV400559 738 bp mRNA EST 05-FEB-2000
LOCUS AV400559 Bombyx mori brain Daizo P0 (just after pupation) Bombyx
DEFINITION AV400559 Bombyx mori brain Daizo P0 (just after pupation) Bombyx
mori cDNA clone br-1449 T3, mRNA sequence.
ACCESSION AV400559
VERSION AV400559.1 GI:6904211
KEYWORDS EST.
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 738)
AUTHORS Mita,K., Morimyo,M., Shinada,T., Okano,K. and Maeda,S.
TITLE Bombyx mori cDNA
JOURNAL Unpublished (2000)
COMMENT Contact: Mita K
Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmits@nirs.go.jp
method:uni-directional, sequence direction:sequenced from T3 primer
(5' -> 3')

Project= Silkorm Genome Program in MAF, and Research for the
Future Program in JSPS. see 'Silkbase',
<http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.
Location/Qualifiers

FEATURES

source
1. 738
/organism="Bombyx mori"
/strain="Daizo"
/db_xref="taxon:7091"
/clone="br-1449"
/clone_id="Bombyx mori brain Daizo P0 (just after
pupation)"
/sex="female/male mixed"
/tissue-type="brain"
/dev_stage="P0 (just after pupation)"
BASE COUNT 255 a 140 c 148 g 195 t
ORIGIN

Query Match 9.18; Score 90.4; DB 29; Length 738;
Best Local Similarity 54.64; Pred. No. 1e-11;
Matches 201; Conservative 0; Mismatches 166; Indels 1; Gaps 1;

OY 522 agatgataatagtaactgagacgaatctatacagatttgaagaacattggaag 581
Db 349 AGACATATTTACACGTCTAAGGAGAGGCTTGTGATATATTACAAAGTATGCCCTAA 408
OY 582 aattttagtggtttaaactaaagtgaaatggaacaaagtgattgcaagtc 641
Db 409 AATCATTTGACAGAGAAAGCTAATTTGAAGAGCTGGCAAGTACATATTTCTGATTTCATC 468
OY 642 acatgacaagagtggtttagtggttttttagcaaaagcaaatgcaaccacattgca 701
Db 469 AGCAAAACGTCATGCGGATTTTGGCTTTCACAGAAAGTCAAAATGCAAGAACCATTTGA 528
OY 702 agaagccgcatcaatttgcattgttgagaagtcagtaacataattgaaacaaatagtg 761
Db 529 AGACGTGCAATTTCTCTTGAATAATAGACAAATTGAGCAACCACTTACACTGACTCGG 588
OY 762 tgtccatatactccaaagaacagataaatacaagattggaagttgataaataatgaa 821
Db 589 CATTCATATCTTTTGTGAACCTGCTTAAAGATTTGTTTTCACAG-TTAATATCTACTTTAAA 647
OY 822 ataataagagacaagtgatagatttgcataccaaaagaagcgatggtcacaagaatcg 881


```

DB 648 AAGTGAACAAATATATGAACTTATCTACCCACAATATCGCAAAATACAAATCAAGTAT 707
OY 882 aaactgt 889
DB 708 AGATCTTT 715

RESULT 13
AU003598
LOCUS AU003598 748 bp mRNA EST 19-JAN-1999
DEFINITION AU003598 Bombyx mori p50(Daizo) Bombyx mori cDNA clone w500315,
mRNA sequence.
ACCESSION AU003598
VERSION AU003598
KEYWORDS EST. GI:4160969
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
; Bombycoidea; Bombycidae; Bombyx.
REFERENCE 1 (bases 1 to 748)
Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
Establishment of cDNA database of Bombyx mori.
Unpublished (1999)
JOURNAL Contact: Mita K
COMMENT Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmilaenr.90.jp
PROJECT = 'CREST project by JST'.

FEATURES
source 1..748
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="w500315"
/clone_1lb="Bombyx mori p50(Daizo)"
BASE COUNT 257 a 144 c 149 g 198 t
ORIGIN
Query Match 9.1%; Score 90.4; DB 106; Length 748;
Best Local Similarity 54.6%; Pred. No. 1e-11;
Matches 201; Conservative 0; Mismatches 166; Indels 1; Gaps 1;

OY 522 agatgtataagtagacagagaacatctatacagatatgaagaacatttgaag 581
DB 329 AGAGCATATTACACGTACTAGAGGAGGAGACCTCTTGATATATTACAGAGTATCGCGGTAA 388
OY 582 aatattgagtggtgaagtttaactaagtgtaattggaatcaccgaagaagtgcagctc 641
DB 389 AATCATTTGACAGAGAACGTAAATTGGAAGAGCTGCGCAATACATATTCGATTGTTTCATC 448
OY 642 acatgacagaagtggtgattaggglltttagcaaaaggaacaatgcaaccacattcga 701
DB 449 ACCAAAACGTGATGGGAGATTGGTCGTTTCAAGAAAGTCAAAATGCCAAGAACCAATTGCA 508
OY 702 agaaagccgattcaatttgcagtgtggaagaatcagtaacaataatgaaccaatagtg 761
DB 509 AGACGTGCGATTTTCTTGAAATAGGACAATTGAGCCCAACAGTTCCACACTGACTCTGG 568
OY 762 tgtccatatccccaagaagacagataaatacagaatatgtgagtttgaagaataatgaa 821
DB 569 CATTTCATATCATTTCTTGAAGCTGCTAAAGATTGTTTACAG-TTAATACACTTCTTAAAA 627
OY 822 ataaatagagacaagtgctatagatttggtaacaaaagaagtgctcacaaaagtcg 881
DB 628 AAGTGAACAAATATATCAAACTTATATACCAACATATCGCAAAATACAAATCAAGTAT 687
OY 882 aaactgt 889
DB 688 AGATCTTT 695

```

```

RESULT 14
AU000274
LOCUS AU000274 660 bp mRNA EST 15-JAN-1999
DEFINITION AU000274 Bombyx mori p50(Daizo) Bombyx mori cDNA clone e40334, mRNA
sequence.
ACCESSION AU000274
VERSION AU000274
KEYWORDS EST. GI:4156518
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia
; Bombycoidea; Bombycidae; Bombyx.
REFERENCE 1 (bases 1 to 660)
Mita,K., Morimyo,M., Shimada,T., Okano,K. and Maeda,S.
Establishment of cDNA database of Bombyx mori.
Unpublished (1999)
JOURNAL Contact: Mita K
COMMENT Genome Research Group
National Institute of Radiological Sciences
Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
Email: kmilaenr.90.jp
PROJECT = 'CREST project by JST'.

FEATURES
source 1..660
/organism="Bombyx mori"
/strain="p50(Daizo)"
/db_xref="taxon:7091"
/clone="e40334"
/clone_1lb="Bombyx mori p50(Daizo)"
BASE COUNT 233 a 127 c 136 g 164 t
ORIGIN
Query Match 9.1%; Score 89.8; DB 106; Length 660;
Best Local Similarity 57.2%; Pred. No. 1.4e-11;
Matches 163; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

OY 522 agatgtataagtagacacagagaacatctatacagatatgaagaacatttgaag 581
DB 308 AGAGCATATTACACGTACTAGAGGAGGAGACCTCTTGATATATTACAGAGTATCGCGGTAA 367
OY 582 aatattgagtggtgaagtttaactaagtgtaattggaatcaccgaagaagtgcagctc 641
DB 368 AATCATTTGACAGAGAACGTAAATTGGAAGAGCTGCGCAATACATATTCGATTGTTTCATC 427
OY 642 acatgacagaagtggtgattaggglltttagcaaaaggaacaatgcaaccacattcga 701
DB 428 ACCAAAACGTGATGGGAGATTGGTCGTTTCAAGAAAGTCAAAATGCCAAGAACCAATTGGA 487
OY 702 agaaagccgattcaatttgcagtgtggaagaatcagtaacaataatgaaccaatagtg 761
DB 488 AGACGTGCGATTTTCTTGAAATAGGACAATTGAGCAACCAAGTTCCACACTGCTGG 547
OY 762 tgtccatatccccaagaagacagataaatacagaatatgtgagtttgaagaataatgaa 806
DB 548 CATTTCATATCATTTCTTGAAGCTGCCCTAAAGATTGTTTTCACAGTT 592

RESULT 15
AU003857
LOCUS AU003857 722 bp mRNA EST 19-JAN-1999
DEFINITION AU003857 Bombyx mori p50(Daizo) Bombyx mori cDNA clone w500674,
mRNA sequence.
ACCESSION AU003857
VERSION AU003857
KEYWORDS EST. GI:4161228
SOURCE domestic silkworm.
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia

```

```
REFERENCE      ; Bombycoidea; Bombycidae; Bombyx.
AUTHORS        Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.
TITLE          Establishment of cDNA database of Bombyx mori
JOURNAL        Unpublished (1999)
COMMENT        Contact: Mita K
                Genome Research Group
                National Institute of Radiological Sciences
                Anagawa 4-9-1, Inage, Chiba 263-8555, Japan
                Email: kmita@nirs.go.jp
                PROJECT = 'CREST project by JST'.
FEATURES
  source        1..722
                /organism="Bombyx mori"
                /strain="p50(Daizo)"
                /db_xref="taxon:7091"
                /clone="w80674"
                /clone_1bp="Bombyx mori p50(Daizo)"
BASE COUNT     246 a      141 c      145 g      190 t
ORIGIN

Query Match
Best Local Similarity 9.1%; Score 89.8; DB 106; Length 722;
Matches 163; Conservative 0; Mismatches 122; Indels 0; Gaps 0;

QY 522 agatggtataagtagaactagaacgaactctatacagatattgaagaacattggaag 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 AGACGATATTACAGCTACTAGAGGAGGAGCTCTGATATATTACAGAGTATCGCCCTAA 377

QY 582 aatattgagtgtaggttaactaagtgaattggcaaataccgaagtgaattgcagctc 641
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 AATCATTTGACAGAGAGAGCTAAATTGGAAGAGCTGGCAGTACATATTCTGATTGTTCAATC 437

QY 642 acatgacagagtgtaggttaggttttttagcaagacaatgcaaccaccattcga 701
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 AGCAAAACGTCATGGGATTTGGGTCGTTTCAAGAAAGGTCAAATGCGAAGACCATTTGA 497

QY 702 agaagccgcatcattgcatgtltgagaagtcagtaacataattgaaaccaatagtcg 761
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 AGACGTCGCTATTTCCTGGAATAATAGGACAAATGAGCCACAGTTCACACTGACTCTGG 557

QY 762 tgtcatatctctcaagaacagagataaatacagatatattgaggt 806
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 CATTCATATCATCTCTAGAACTGCTAAAGATTTGTTTCACAGTT 602
```

Search completed: June 8, 2001, 05:10:35
Job time: 4325 sec

QY	63	NEDGQVRVSHLLTIKNNQSPKSPDIPRDRSDISIOILKHLIERITISGEVKLSLEAN	122
Db	63	SEEGNIRCHSHLLVKNHDSRRPSSWRAE-IIRTKKEAREILKHQHOPRINIGETRLDLAM	121
QY	123	TESDCSSHDGRGDLGFFSGKQOMOPPEEAFNLIHGEVSNIIETNSGVHILOR	175
Db	122	SESDCSSARKKGDGLGFFGREGEMQKEFEFAFALOPQOVSDIYVSGSGHLIER	174
RESULT	2		
ID	060045	PRELIMINARY:	PRT; 182 AA.
AC	060045;		
DT	01-AUG-1998	(TREMBLrel. 07, Created)	
DT	01-AUG-1998	(TREMBLrel. 07, Last sequence update)	
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)	
DE	PEPTIDYLPROLYL ISOMERASE (EC 5.2.1.8).		
GN	SSPI.		
OS	Neurospora crassa.		
OC	Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.		
CC	NCBI_TaxID=5141;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Kops O., Tropschug M.,		
RL	Submitted (May-1998) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AJ006023; CAA06818.1; --		
DR	HSSP; Q13526; IPIN.		
DR	INTERPRO: IPR0000297; --		
DR	INTERPRO: IPR001202; --		
DR	INTERPRO: IPR002349; --		
DR	PFAM; PF00397; WW; 1.		
DR	PFAM; PF00639; Rotamase; 1.		
DR	PRINTS; PR00403; WMDOMAIN.		
DR	PROSITE; PS01159; WW_DOMAIN_1; UNKNOWN_1.		
DR	PROSITE; PS0020; WW_DOMAIN_2; 1.		
DR	PRODOM; PD002510; --; 1.		
DR	Isomerase.		
QW	SEQUENCE 182 AA; 20673 MW; AC8D0497DFF7339 CRC64;		
QY	Query Match	45.0%; Score 419.5; DB 3; Length 182;	
Db	Best Local Similarity	44.3%; Pred. No. 1.3e-29;	
QY	Matches 82; Conservative 40; Mismatches 48; Indels 15; Gaps		
QY	1	MAST-STGCLPMMTLRVSRSHNKEYFLNSTRESSMDPYGDKVELNVIKFNKNNGYK	59
Db	1	MSNTIETGLPEMEVRRHOSKMLPYFNNSATITSRKEPSSGDVVDLKIYMAV---HS	56
QY	60	PLVNE-----DGQVRVSHLLTIKNNQSPKSPDGISRTDESIOILKHLERI	110
Db	57	PTSOQOQOQOQOQPGKIRCAHLVKNHDSRRPSSWRESE-IIRTKQEALTLTLOGEQRI	115
QY	111	LSGEVKLSLEANTESDCSSHDGRGDLGFFSGKQOMOPPEEAFNLIHGEVSNIIETNSGV	170
Db	116	KSGSISLDELATSEDCSSARRRGDLGFGKRGDMQKEFDAAFAALPKPEISIDVTSAGL	175
QY	171	HILOR	175
Db	176	HLIER	180
RESULT	3		
ID	074448	PRELIMINARY:	PRT; 175 AA.
AC	074448;		
DT	01-NOV-1998	(TREMBLrel. 08, Created)	
DT	01-NOV-1998	(TREMBLrel. 08, Last sequence update)	
DT	01-OCT-2000	(TREMBLrel. 15, Last annotation update)	
DE	PROBABLE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE (EC 5.2.1.8).		
GN	SPEC1C4.03.		
OS	Schizosaccharomyces pombe (Fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;		

```

OC Schizosaccharomycetaceae; Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA purelle B., Goffeau A., Wood V., Rajandream M.A., Barrell B.G.;
RL Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.
CC -I- CATALYTIC ACTIVITY: CIS-TRANS ISOMERIZATION OF PROLINE IMIDIC
CC PEPTIDE BONDS IN OLIGOPEPTIDES.
CC -I- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -I- SIMILARITY: BELONGS TO THE PPI/C/PAVULIN FAMILY OF ROTAMASES.
DR EMBL: AL031535; CAA20742.1; -.
DR HSSP: Q13526; 1PIN.
DR INTERPRO: IPR001302; -.
DR INTERPRO: IPR000297; -.
DR PFM: PFM0397; MW; 1.
DR PFM: PFM0639; Rotamase; 1.
DR PROSITE: PS50020; MW_DOMAIN_2; 1.
DR PRODOM: PD002510; -. 1.
DR Isomerase; Rotamase; Nuclear protein.
KW DOMAIN 4 38 MW DOMAIN.
FT PIC-LIKE.
SQ SEQUENCE 175 AA; 19773 MW; ABA637835471BD25 CRC64;

Query Match 43.9%; Score 409; DB 3; Length 175;
Best Local Similarity 46.0%; Pred. No. 1e-26;
Matches 80; Conservative 37; Mismatches 53; Indels 4; Gaps 4

OY 4 TSTGLPNNNTIVSSHNKEVFLNOSTNNESSMDPRYGTQKEVLAIAFKNNGYRL-V 62
Db 2 SNTGTPKPIVIVISRSNRPFYFNETHESLMEPPATVMAALKFNAIELGESVPTPEA 61
OY 63 NEDGQVRSVSHLLIKNNQSKPKMSKSPDISRTDPSIOILKKHLERLV-SEGVKSELA 121
Db 62 SWSRPIRASHLLVKNRRESRPSKWE-EHITRSKEEA-RKLEHVEQGLKSSGSVSHMDLA 119
OY 122 NTSDCCSSHDRCGLDGFESKSGOMQDPFEFAFNALHGEVSNITETNSGVHIIOR 175
Db 120 MKESCCSSARRGELGEGFGRDEMOKPFEDAAALAKGEISGVETSSGPHIIR 173

RESULT 4
ID 0919K6 PRELIMINARY; PRT; 159 AA.
AC 0919K6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PROLYL ISOMERASE PIN1.
GN PIN1.
OS Xenopus laevis (African Clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20165035; PubMed=10698738;
RA Winkler K.E., Swenson K.I., Kornbluth S., Means A.R.;
RT "Regulation of the prolyl isomerase pin1 for the replication
RT checkpoint.:"
RL Science 287:1644-1647(2000).
DR EMBL: AF239760; AAF43897.1; -.
KW Isomerase.
SQ SEQUENCE 159 AA; 17671 MW; 4E54F66F16C9DF1C CRC64;

Query Match 40.1%; Score 373.5; DB 13; Length 159;
Best Local Similarity 44.4%; Pred. No. 1.2e-25;
Matches 75; Conservative 32; Mismatches 45; Indels 17; Gaps 3

```

[illegible]

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dudin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.U., Mei M.-H., Ibegwan C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matlet B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA She B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Massarman D.A., Weinstein G.M., Weinsteinbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003568; AAF50829.1; -.
 DR HSSP: Q13526; IPTN.
 DR FLYBASE: FBgn0015379; dod.
 DR INTERPRO: IPR000297; -.
 DR INTERPRO: IPR001202; -.
 DR PFAM: PF00397; Rotamase; 1.
 DR PFAM: PF00639; Rotamase; 1.
 DR PROSITE: PS01096; PRIC_PPIASE; 1.
 DR PROSITE: PS01159; WW_DOMAIN; 1.
 DR PROSITE: PS50020; WW_DOMAIN; 2; 1.
 DR SEQUENCE 166 AA; 18376 MW; 3B4306FA930E7259 CRC64;

Query Match 34.9%; Score 325; DB 5; Length 166;
 Best Local Similarity 41.7%; Pred. No. 2.4e-21;
 Matches 70; Conservative 30; Mismatches 58; Indels 10; Gaps 3;
 QY 8 LPPNNTIVRSRSHNKEFLPNOSTNSSMDPYGTDEKVLNATIAFKNGGYPLVNEGQ 67
 DB 7 LPDGEKRTSRSTGMSXYLYMYTKESQMDQ-TP-----AKKAGGSGAGGADPDE 57
 QY 68 VVSHLLIKNNQSRKPKSKSPGDISRTDESIOILKKHLERILSGEYKLSFLANTESDC 127
 DB 58 VVCHLLIKKHKSRKSRPSKREAN-ITRTKEAQLLEIYRKTIVOQEAFTDELANSYSDC 116
 QY 128 SSHDRGDLGFFSGKQMPPEEAAFNHVGESVNIITNSGVHILORT 175
 DB 117 SSARRGGDLGKFRGQMOAFEDAFKLVNQLSGIVSDSLHILR 164

RESULT 8
 QYLER8 PRELIMINARY; PRT; 118 AA.
 AC QYLER8;
 DT 01-OCT-2000 (TREMBlrel. 15 Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
 GN PARY12.8.
 OS *Digitalis lanata* (Foxglove).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
 CC Lamiales; Scrophulariaceae; Digitalis.
 OX NCBI_TaxID=49450;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=EHK;
 Metzner M., Ruecknagel P., Kuelbert G., Luckner M.;

RT "Isolation and characterization of a new parvulin like peptidyl-prolyl
 RT cis/trans isomerase from *Digitalis lanata*.";
 RT Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AJ133755; CAB94994.1; -.
 KW Isomerase.
 SQ SEQUENCE 118 AA; 12834 MW; 131B74FB4AC3F229 CRC64;

Query Match 32.8%; Score 305.5; DB 10; Length 118;
 Best Local Similarity 52.6%; Pred. No. 8.2e-20;
 Matches 60; Conservative 21; Mismatches 30; Indels 3; Gaps 1;
 QY 67 QVRVSHLLIKNNQSRKPKSKSPDG--ISRTDESIOILKKHLERILSGEYKLSFLANT 123
 DB 5 KVRASHLLIKHGGSRKSKSMKDPDLSLATRRDDAVSOLSLROELSDPASFDLASR 64
 QY 124 ESDCSSHDRGDLGFFSGKQMPPEEAAFNHVGESVNIITNSGVHILORT 177
 DB 65 HSHCSSARRGGDLGFRGQMPPEEATFALKVGEISIDIVDTSGVHIKRTG 118

RESULT 9
 QYLER8 PRELIMINARY; PRT; 119 AA.
 AC QYLER8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PUTATIVE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
 GN AT2G18040.
 OS *Arabidopsis thaliana* (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 CC Brassicales; Brassicaceae; Arabidopses.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffatt K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carreira A.U., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nielsen M.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome II of *Arabidopsis thaliana*.";
 RL Nature 402:761-768(1999).
 DR EMBL: AC006201; AAD20122.1; -.
 DR HSSP: Q13526; IPTN.
 DR INTERPRO: IPR000297; -.
 DR PFAM: PF00639; Rotamase; 1.
 DR PROSITE: PS01096; PRIC_PPIASE; 1.
 KW Isomerase.
 SQ SEQUENCE 119 AA; 13015 MW; E926CB566E76A0A3 CRC64;

Query Match 32.6%; Score 303.5; DB 10; Length 119;
 Best Local Similarity 51.3%; Pred. No. 1.2e-19;
 Matches 58; Conservative 24; Mismatches 28; Indels 3; Gaps 1;
 QY 67 QVRVSHLLIKNNQSRKPKSKSPDG--ISRTDESIOILKKHLERILSGEYKLSFLANT 123
 DB 6 QVKAHHLLIKHGGSRKSKSMKDPDKILITTRAAVQLKSRIDIVSGKANFEVATR 65
 QY 124 ESDCSSHDRGDLGFFSGKQMPPEEAAFNHVGESVNIITNSGVHILORT 176
 DB 66 VSDCSSARRGGDLGFRGQMPPEEATYALKVDISPIDVTGSGVHIKRT 118
 RESULT 10
 P90527 PRELIMINARY; PRT; 243 AA.
 AC P90527;
 P90527;

DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
GN PINA. (FRAGMENT).
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Dictyostelida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AX4;
RA Loomis W.F., Iranfar N.;
RL Submitted (Nov-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78757; AAB36960.1; -.
DR HSSP; Q13526; IPIN.
DR INTERPRO; IPR000253; -.
DR INTERPRO; IPR000297; -.
DR PFAM; PF00498; FHA; 1.
DR PFAM; PF00639; Rotamase; 1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
DR NON_TER 1
SQ SEQUENCE 243 AA; 26794 MW; 1596001BB784500A CRC64;

Query Match 29.1%; Score 271.5; DB 5; Length 243;
Best Local Similarity 37.4%; Pred. No. 2e-16;
Matches 65; Conservative 29; Mismatches 57; Indels 23; Gaps 4;

QY 4 TSGTLPNNNTIRSRSHNKEYFL-NOSTNESSWDPPYGTGDEKVLNAYIKFKNNGYKPL 61
DB 89 TPTTVENETTFKFG-SSSKHFIKGTNTNPSS-----SSSSSSSS 128
QY 62 VNEDGQVRVSHLLIKNNOSRPKSKSPDGISRTDESIOILKKHLRIISGEVKLSLA 121
DB 129 SSPEKPVYTCRHLLVKNQGSNPFSSWRE-SKITTKERAIKLEKRYRTISGSAITPEDLA 187
QY 122 NTESDSSHDRCGLDGFPSKGQMPPEEAFFNLHVGESNIETNSGVHILOR 175
DB 188 HKNSDCSSAKRGGLDLPFRKGQMRPEFDCAFSLKGVESGIYDFTDSGVHIER 241

RESULT 11
Q42334 PRELIMINARY; PRT; 69 AA.
AC Q42334;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, last annotation update)
DE PEPTIDYL PROLYL CIS-TRANS ISOMERASE (FRAGMENT).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-DRY SEEDS OF A. THALIANA ECOTYPE COLUMBIA;
RA Cooke R., Laudie M., Raynal M., Delseny M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; F13919; CAA23077.1; -.
DR HSSP; Q13526; IPIN.
DR MENDEL; 12876; Archt; 2076; 12876.
DR INTERPRO; IPR000297; -.
DR PFAM; PF00639; Rotamase; 1.
DR PROSITE; PS01096; PPI_CPIINASE; 1.
KW Isomerase.
FT NON_TER 1
SQ SEQUENCE 69 AA; 7333 MW; 03C2DF735ED6F253 CRC64;

Query Match 22.6%; Score 211; DB 10; Length 69;
Best Local Similarity 58.2%; Pred. No. 8.9e-12;
Matches 39; Conservative 13; Mismatches 15; Indels 0; Gaps 0;

QY 110 ILSGEVLSELANTESDSSHDRCGLDGFPSKGQMPPEEAFFNLHVGESNIETNSG 169
DB 2 IVSGKANFEVATRVSDCSSAKRGGLDGSFRGOMQMPFEATYALKVDISIDVDSG 61
QY 170 VHIQRT 176
DB 62 VHIKRT 68

RESULT 12
Q9KDN4 PRELIMINARY; PRT; 333 AA.
AC Q9KDN4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE PROTEIN SECRETION (POST-TRANSLATION CHAPERONIN).
GN PRSA.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=8665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA Takami H., Nakasone K., Takaki Y.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF001511; BAB04896.1; -.
SQ SEQUENCE 333 AA; 37266 MW; A0B77383BC1AD61E CRC64;

Query Match 17.9%; Score 166.5; DB 2; Length 333;
Best Local Similarity 30.1%; Pred. No. 5.5e-07;
Matches 50; Conservative 25; Mismatches 40; Indels 51; Gaps 7;

QY 21 NKEYFLNOSTNESSWDPPYGTDR-----EVLNAVYAKERNNGYKPLVNEGG--QY 68
DB 114 NKEEFVNEYIL-----PVLVEKLAIVGEVETDEKQAYPDE-----NRDSLIEV 158
QY 69 RVSHLLIKNNOSRPKSKSPDGISRTDESIOILKKHLRIISGEVKLSLA 127
DB 159 EASHILVEDET-----AEEVLDRLAGD-DFEELASEVSYDP 195
QY 128 SSHDRGDLGFFSGQMPPEEAFFNLHVGESNIETNSGVHIL 173
DB 196 SAAANNCDLGFSGKGMVPEFEAAFMETIDVSEPEVSTGYGHIIT 241

RESULT 13
Q74049 PRELIMINARY; PRT; 92 AA.
AC Q74049;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, last annotation update)
DE PEPTIDYL-PROLYL CIS-TRANS ISOMERASE.
GN PINA.
OS Cenarchaeum symbiosum.
OC Archaea; Crenarchaeota; Cenarchaeum.
OX NCBI_TaxID=46770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A;
RX MEDLINE=98422450; PubMed=9748430;
RA Schleper C., DeLong E.F., Preston C.M., Feldman R.A., Wu K.Y.,
RA Swanson R.V.;
RT "Genomic analysis reveals chromosomal variation in natural populations
of the uncultured psychrophilic archaeon Cenarchaeum symbiosum.";
DR J. Bacteriol. 180:5003-5009(1998).
DR EMBL; AF083071; AAC62692.1; -.
DR HSSP; Q13526; IPIN.
DR INTERPRO; IPR000297; -.

```
DR PFAM: PF00639; Rotamase; 1.
DR PRODOM: PD002510; -; 1.
KW Isomerase.
SQ SEQUENCE 92 AA; 10067 MW; AADAB5FE60C518FE CRC64;

Query Match
Best Local Similarity 33.9%; Score 157; DB 1; Length 92;
Matches 38; Conservative 22; Mismatches 28; Indels 24; Gaps 4;

QY 67 QVRVSHLLIKNNOSRRKPKSWKSPDGISRTDSEIQILKKHLERILSGEYKLSLANTESD 125
DB 4 KIKCSHLLVKK-----QGEALAV---QERLRAGE-KFGKIAKELSI 40

QY 126 DCSHSDRGDDLGFESKGMQOPPEEAFAFNLHGEVSNIIETNSGVHILQRTG 177
DB 41 DGSAAKRDGSLGFGKMKVKKPFEDAAFRLLQGEVSEPVKSEFGVYKIRLG 92

RESULT 14
023727 PRELIMINARY; PRT; 221 AA.
AC 023727;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PEPIDYL-PROLYL CIS-TRANS ISOMERASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Brassicales; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC NCBI_TaxID=3702;
OX [1]
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T-DNA TAGGED LINE SK3-1; TISSUE=LEAVES;
RA Babyichuk E., Funagthong M., Inze D., van Montagu M., Kushnir S.;
RU Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
DR EMBL: Z86095; CAB06699.1; -.
DR HSSP: Q13526; IPIN.
DR INTERPRO: IPR000297; -.
DR INTERPRO: IPR001763; -.
DR PFAM: PF00581; Rhodanese; 1.
DR PFAM: PF00639; Rotamase; 1.
DR PRODOM: PD002510; -; 1.
KW Isomerase.
SQ SEQUENCE 221 AA; 24744 MW; 491BE5C4C5E6117 CRC64;

Query Match
Best Local Similarity 15.2%; Score 142; DB 10; Length 221;
Matches 36; Conservative 17; Mismatches 33; Indels 22; Gaps 3;

QY 67 QVRVSHLLIKNNOSRRKPKSWKSPDGISRTDSEIQILKKHLERILSGEYKLSLANTESD 126
DB 17 EILVGHLLVKNK-----DVLEFAELQKKFLDGE-EMSDLALEYSI 55

QY 127 CSHSDRGDDLGFESKGMQOPPEEAFAFNLHGEVSNIIETNSGVHILQ 174
DB 56 CPSKRDGILGKGMVPEEEAFAFKAELNQVVR-CRTQFGLHILQ 102

RESULT 15
093SP0 PRELIMINARY; PRT; 347 AA.
AC 093SP0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PUTATIVE ROTAMASE.
GN NMA2206.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed-10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Fellwell T., Hamlin N., Holroyd S.,
RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162758; CAB85417.1; -.
SQ SEQUENCE 347 AA; 37126 MW; A387AD675F86FA51 CRC64;

Query Match
Best Local Similarity 13.1%; Score 122; DB 2; Length 347;
Matches 35; Conservative 14; Mismatches 41; Indels 22; Gaps 2;

QY 67 QVRVSHLLIKNNOSRRKPKSWKSPDGISRTDSEIQILKKHLERILSGEYK----LSELAN 122
DB 207 QYRAQHLLKAD-----SENAVGAESTRKITYGEARSGTDFSSILAR 248

QY 123 TESDCSHDRGDDLGFESKGMQOPPEEAFAFNLHGEVSNIIETNSGVHILQ 174
DB 249 QVSODASAGNGDDLGFADGVWVPAPFEAAVHALKPGVGVAPVPTQFGWHIIR 300
```

Search completed: June 8, 2001, 09:38:38
Job time: 48 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2001, 05:10:35 ; Search time 1747.82 Seconds
(without alignments)
99.966 Million cell updates/sec

Title: us-09-507-242-3

Perfect score: 20
Sequence: 1 ccagatgataaagtagaac 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9623517 segs, 4368049070 residues

Total number of hits satisfying chosen parameters: 19247034

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

```

```

190: em_gss_pln1:*
191: em_gss_pln2:*
192: em_gss_pro:*
193: em_gss_rod1:*
194: em_gss_rod2:*
195: em_gss_rod3:*
196: em_gss_rod4:*
197: em_gss_rod5:*
198: em_gss_vrt1:*
199: em_gss_vrt2:*
200: em_gss_vrt3:*
201: gb_gss1:*
202: gb_gss2:*
203: gb_gss3:*
204: gb_gss4:*
205: gb_gss5:*
206: gb_gss6:*
207: gb_gss7:*
208: gb_gss8:*
209: gb_gss9:*
210: gb_gss10:*
211: gb_gss11:*
212: gb_gss12:*
213: gb_gss13:*
214: gb_gss14:*
215: gb_gss15:*
216: gb_gss16:*
217: gb_gss17:*
218: gb_gss18:*
219: gb_gss19:*
220: gb_gss20:*
221: gb_gss21:*
222: gb_gss22:*
223: gb_gss23:*
224: gb_gss24:*
225: gb_gss25:*
226: gb_gss26:*
227: gb_gss27:*
228: gb_gss28:*
229: gb_gss29:*
230: gb_gss30:*
231: gb_gss31:*
232: gb_gss32:*
233: gb_gss33:*
234: gb_gss34:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	16.8	84.0	225	127	BB208038
2	16.8	84.0	611	31	AV655972
3	16.8	84.0	630	214	AO980511
C 4	16.8	84.0	710	16	AI132934
5	16.8	84.0	846	149	BF540037
C 6	16.8	84.0	894	151	BF692672
C 7	16.4	82.0	188	139	BE807035
8	16.4	82.0	213	16	AI120952
9	16.4	82.0	222	11	AA717739
10	16.4	82.0	242	7	AA458264
11	16.4	82.0	250	149	BF542933
C 12	16.4	82.0	271	169	BF812297
13	16.4	82.0	281	152	F06213
14	16.4	82.0	291	152	F06677
15	16.4	82.0	298	113	AW313588
16	16.4	82.0	308	152	F06211
17	16.4	82.0	320	155	R22624
18	16.4	82.0	331	113	AW313587

Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
location/Qualifiers
1. 611

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="GICEMD12"
/clone_1lb="GIC"
/tissue_type="corresponding non cancerous liver tissue"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: p Bluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 253 a 104 c 80 g 174 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 31; Length 611;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ccagatggtataagtagaac 20
||| ||||| ||||| |||||
Db 445 CCAATGATATCAGTAGAAC 464

RESULT 3
LOCUS A0980511 630 bp DNA 29-JAN-2000
DEFINITION RPCI-23-350G7.TU RPCI-23 Mus musculus genomic clone RPCI-23-350G7,
DNA sequence.
ACCESSION A0980511
VERSION A0980511.1 GI:6812812
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 630)
Zhao S., Nieman, W., Feldblum, T., Malek, J., Shatsman, S., Akinet
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.
and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

TITLE
JOURNAL
COMMENT
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 350 row: G column: 7
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. 630

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-350G7"
/clone_1lb="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or

brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBAC3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 210 a 148 c 125 g 147 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 214; Length 630;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ccagatggtataagtagaac 20
||| ||||| ||||| |||||
Db 148 CCACATGATATACAGAGAAC 167

RESULT 4
LOCUS A1132934/c 710 bp mRNA EST 11-NOV-1999
DEFINITION HAI517 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
sequence.
ACCESSION A1132934
VERSION A1132934.1 GI:6360250
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 710)
Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
Expression profile analysis of a human fetal liver cDNA library
Unpublished (1998)
Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyt48@yahoo.com

TITLE
JOURNAL
COMMENT
Expression profile analysis of a human fetal liver cDNA library
Unpublished (1998)
Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yyt48@yahoo.com

FEATURES
source
Location/Qualifiers
1. 710

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lb="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="Vector: pCDNA1"

BASE COUNT 210 a 100 c 118 g 280 t 2 others
ORIGIN

Query Match 84.0%; Score 16.8; DB 16; Length 710;
Best Local Similarity 90.0%; Pred. No. 2.3e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 ccagatggtataagtagaac 20
||| ||||| ||||| |||||
Db 140 CCAATGATATCAGTAGAAC 121

RESULT 5
LOCUS BF540037 846 bp mRNA EST 11-DEC-2000
DEFINITION 602050669F1 NCL_CGAP_SG2 Mus musculus cDNA clone IMAGE:418937 5',
mRNA sequence.
ACCESSION BF540037
VERSION BF540037.1 GI:11627418
KEYWORDS EST.
SOURCE house mouse.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 846)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM9516 row: 1 column: 18
High quality sequence stop: 107.
Location/Qualifiers

FEATURES

source

1. 846
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4189937"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: salivary gland. Vector: pCW-SPORE6; Site: 1;
Nori; Site: 2; Salt; cloned unidirectionally. Primer: Oligo
dT. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 302 a 100 c 191 g 253 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 149; Length 846;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
1 |||||
Db 735 CGAGATTGTAATGAGAAC 754

RESULT 6
BF692672/c 894 bp mRNA EST 22-DEC-2000
LOCUS 602249045f1 NIH-MGC-62 Homo sapiens cDNA clone IMAGE:434291 5',
DEFINITION mRNA sequence.
ACCESSION BF692672
VERSION BF692672.1 GI:11978080
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 894)
NIH-MGC <http://mhc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: ATCC/DCMD/DRP
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LLM1203 row: 1 column: 12
High quality sequence stop: 606.
Location/Qualifiers

FEATURES

source

1. 894
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:434291"
/clone_lib="NIH-MGC-62"
/tissue_type="melanotic melanoma, high MDR"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin. Vector: pDNR-LIB (Clontech); Site: 1;
SfiI (ggccgctggcc); Site 2: SfiI (ggccataggcc);
double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
sequence: 5'-ATCTAGAGCGCGCGCGCGCATG-dT(30)BN-3'
(where B = A, C, G or T). Average
insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 214 a 215 c 213 g 252 t
ORIGIN

Query Match 84.0%; Score 16.8; DB 151; Length 894;
Best Local Similarity 90.0%; Pred. No. 2.4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ccagatggtataagtagaac 20
1 |||||
Db 648 CGAGATTGTAATGAGAAC 629

RESULT 7
BE807035/c 188 bp mRNA EST 21-NOV-2000
LOCUS BE807035
DEFINITION Gm-c1047-2584 5', mRNA sequence.
ACCESSION BE807035
VERSION BE807035.1 GI:10238147
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
Fabales; Fabaceae; Papilionoideae; Glycyne.
1 (bases 1 to 188)
Shoemaker R., Kelm P., Vodkin L., Eprelding J., Corvett V., Khanna
, A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J., Beck C.,
Wylie T., Underwood K., Steptoe M., Theising B., Allen M., Bowers
, Y., Person B., Swaller T., Gibbons M., Pape D., Harvey N., Schurk
, R., Ralster E., Kohn S., Shin T., Jackson Y., Cardenas M., McCann
, R., Waterston R. and Wilson R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R./Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 659 Std Error: 0.00
High quality sequence stop: 174.
Location/Qualifiers

FEATURES

source

1. 188
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1047-2584"
/clone_lib="Gm-c1047"

BASE COUNT 83 a 56 c 60 g 51 t
ORIGIN 1996)"

Query Match 82.0%; Score 16.4; DB 149; Length 250;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 agatcgataagtagaac 20
|||||
Db 168 AGATGCTATAATAGAAC 185

RESULT 12
BF812297 271 bp mRNA EST 12-JAN-2001
LOCUS
DEFINITION RC2-C10090-211100-017-b01 C10090 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF812297
VERSION BF812297.1 GI:12141908
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 271)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Negai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202653
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC2&tl=RC2-C10090-
211100-017-b01&f3=2000-11-21&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 26
Location/Qualifiers
1. .271
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="C10090"
/dev_stage="Adult"
/note="Organ: colon; ins: Vector: puc18; Site:1: Smal;
Site:2: Smal; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

BASE COUNT 72 a 60 c 49 g 90 t
ORIGIN

Query Match 82.0%; Score 16.4; DB 169; Length 271;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 agatcgataagtagaac 20

Db 128 AGATGCTATAATAGAAC 111
|||||

RESULT 13
F06213 281 bp mRNA EST 19-FEB-1995
LOCUS
DEFINITION HSC10G011 normalized infant brain cDNA Homo sapiens cDNA clone
c-10g11, mRNA sequence.
ACCESSION F06213
VERSION F06213.1 GI:6770029
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 281)
AUTHORS Aulfray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houlgatte,R., Juneau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
Sebastiani-Kabakchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
COMMENT Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1,rue de l'interationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33166472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress library: idt: C; Genexpress_sequence_idt: ylc-10g11
Seq primer: (-21)M13 universal.
Location/Qualifiers
1. .281
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="c-10g11"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site:1: HindIII;
Site:2: NotI; sex:Female; dev_stage=3 months old;
isolate-muscular atrophy patient; tissue-type-total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B. Soares, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 92 a 52 c 65 g 70 t 2 others
ORIGIN

Query Match 82.0%; Score 16.4; DB 152; Length 281;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 agatcgataagtagaac 20
|||||
Db 175 AGATGCTATAATAGAAC 192

RESULT 14
F06677 291 bp mRNA EST 20-FEB-1995
LOCUS
DEFINITION HSC1G0091 normalized infant brain cDNA Homo sapiens cDNA clone
c-19g09, mRNA sequence.
ACCESSION F06677
VERSION F06677.1 GI:672280
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 291)
AUTHORS Auffrey,C., Behar,G., Bois,F., Bouchier,C., da Silva,C., Devignes
,M.D., Duprat,S., Houigatte,R., Jumeau,M.N., Lamy,B., Lorenzo,F.,
Mitchell,H., Mariage-Samson,R., Pletu,G., Pouliot,Y.,
Sebastiani-Kabakchis,C. and Tessier,A.
IMAGE: molecular integration of the analysis of the human genome
and its expression
C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
95277534
JOURNAL MEDLINE
COMMENT Contact: Genethon
Genexpress-Genethon
Genethon Centre de recherche sur le Genome Humain
1, rue de l'Internationale, BP60 91002 EVRY Cedex, FRANCE
Tel: 33169472800
Fax: 33160778698
Email: genexpress@genethon.fr
Single read.
Genexpress_library_id: C: Genexpress_sequence_id: y1c-1g909
Seq primer: (-21)M13-universal.
Location/Qualifiers
1..291
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="normalized infant brain cDNA"
/sex="Female"
/tissue_type="total brain"
/dev_stage="3 months old"
/note="Organ: brain; Vector: lafmid BA; Site.1: HindIII;
Site.2: NotI; sex=Female; dev_stage=3 months old;
Isolate=muscular atrophy patient; tissue_type=total brain
; total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the
lafmid BA vector. Clone library from B.Saeres, Psychiatry
Dept. Columbia University, USA. Normalization_method:
Bento Soares, P.N.A.S in press"

BASE COUNT 94 a 53 c 68 g 72 t 4 others
ORIGIN

Query Match 82.0%; Score 16.4; DB 152; Length 291;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 agatgtataagtagaac 20
|||||
DB 175 AGATGCTATTAATAGAAC 192

RESULT 15
AW313588 298 bp mRNA EST 08-JAN-2001
LOCUS 8892 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
DEFINITION AW313588
ACCESSION AW313588
VERSION AW313588.1 GI:6742808
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 298)
AUTHORS Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M.,
Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid
,W.W. and Keele,J.W.
Design and use of four pooled tissue normalized cDNA libraries for
EST discovery in cattle
Unpublished (2000)
JOURNAL COMMENT Contact: Smith TP

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACAP
BACKWARD: GTTTCCACGACGACGAG
Plate: 142 row: K column: 5
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1..298
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site.1: XbaI; Site.2: XhoI;
library made from pooled tissue from testis, thymus,
semiteendonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

BASE COUNT 98 a 55 c 72 g 73 t
ORIGIN

Query Match 82.0%; Score 16.4; DB 113; Length 298;
Best Local Similarity 94.4%; Pred. No. 3.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 agatgtataagtagaac 20
|||||
DB 175 AGATGCTATTAATAGAAC 192

Search completed: June 8, 2001, 05:10:38
Job time: 4328 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2001, 06:02:45 ; Search time 1501.23 seconds
(without alignments)
196.477 Million cell updates/sec

Title: US-09-507-242-6
Perfect score: 20
Sequence: 1 caatgacgggaacgttcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1283235 segs, 7373929652 residues

Total number of hits satisfying chosen parameters: 2566470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenDbml:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pac1:*
10: gb_pac2:*
11: gb_ph:*
12: gb_p11:*
13: gb_p12:*
14: gb_p13:*
15: gb_p14:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rnd:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rnd:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_r01:*
95: gb_r02:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.8	84.0	980	9 AX065753	AX065753 Sequence
2	16.8	84.0	1404	9 AX065751	AX065751 Sequence
3	16.8	84.0	4161	2 BACGIDA	M65289 Bacillus st
4	16	80.0	674	1 AF083905	AF083905 Salmonell
5	16	80.0	674	1 AF083912	AF083912 Salmonell
6	16	80.0	12099	1 AE000941	AE000941 Methanoba
7	16	80.0	44237	6 CELY65B4BR	AC024847 Caenorhab
8	16	80.0	267118	60 AC006889	AC006889 Caenorhab
9	15.8	79.0	14625	1 AE005077	AE005077 Halobacte
10	15.8	79.0	22440	3 MTV006	AL021006 Mycobacte
11	15.8	79.0	22480	5 CELF44E7	AF016421 Caenorhab

12	15.8	79.0	31205	5	CELC06G1	U41014 Caenorhabdi
C 13	15.8	79.0	47436	67	AC087901	AC087901 Mus muscu
C 14	15.8	79.0	72593	78	AC021519	AC021519 Homo sapi
C 15	15.8	79.0	110000	84	LMFLCHR34_1	Continuation (2 of
C 16	15.8	79.0	174801	64	AC014376	AC014376 Drosophila
C 17	15.8	79.0	174823	61	AC009808	AC009808 Homo sapi
C 18	15.8	79.0	176340	78	AF216671	AF216671 Homo sapi
C 19	15.8	79.0	179989	64	AC016451	AC016451 Homo sapi
C 20	15.8	79.0	285978	4	AE003551	AE003551 Drosophila
C 21	15.8	79.0	303650	4	AP001519	AP001519 Bacillus
C 22	15.4	77.0	61920	66	AC021270	AC021270 Homo sapi
C 23	15.4	77.0	82315	12	AB006707	AB006707 Arabidops
C 24	15.4	77.0	100727	63	AC013831	AC013831 Drosophila
C 25	15.4	77.0	104204	13	ATT9C5	AL132964 Arabidops
C 26	15.4	77.0	107931	12	AC007478	AC007478 Arabidops
C 27	15.4	77.0	133195	12	AC007123	AC007123 Arabidops
C 28	15.4	77.0	134610	76	AC079302	AC079302 Homo sapi
C 29	15.4	77.0	194041	76	AC079349	AC079349 Homo sapi
C 30	15.4	77.0	349061	3	NMA22491	AL167353 Neisseria
C 31	15.2	76.0	126	54	G54042	G54042 1562 Human
C 32	15.2	76.0	490	3	UBA244349	AJ244349 unculture
C 33	15.2	76.0	490	3	UBA244355	U09529 Salmonella
C 34	15.2	76.0	1030	3	STU09529	AJ253628 Agriicultu
C 35	15.2	76.0	1073	2	ASO252625	AJ2020540 Mus muscu
C 36	15.2	76.0	1103	94	AF020540	AK025868 Homo sapi
C 37	15.2	76.0	1705	89	AK025868	AJ250022 Mycobacte
C 38	15.2	76.0	1856	92	MAY250022	AL137722 Homo sapi
C 39	15.2	76.0	2856	92	HSMB02238	K01996 E.coli xyla
C 40	15.2	76.0	4176	2	ECOXYLABA	X04691 E.coli gen
C 41	15.2	76.0	4176	2	ECOXYLABA	M17381 Streptococc
C 42	15.2	76.0	4995	3	STRTFT	E03388 Xylose isom
C 43	15.2	76.0	7801	10	E03388	AF098506 Caenorhab
C 44	15.2	76.0	8770	5	CELH29C22	Z68886 Human DNA s
C 45	15.2	76.0	10589	92	HSI21F12C	

ALIGNMENTS

RESULT 1						
AX065753	980 bp	DNA				
DEFINITION	Sequence 97 from Patent WO0100842.	PAT	24-JAN-2001			
ACCESSION	AX065753					
VERSION	AX065753.1	GI:12543465				
KEYWORDS						
SOURCE						
ORGANISM	Corynebacterium glutamicum.					
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;					
	Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;					
	Corynebacterium.					
	1 (bases 1 to 980)					
REFERENCE	Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O. and Habermayer, G.					
AUTHORS	Corynebacterium glutamicum genes encoding proteins involved in					
TITLE	homeostasis and adaptation					
JOURNAL	Patent: WO 0100842-A 97 04-JAN-2001;					
	Basf AKTIENGESellschaft (DE)					
FEATURES	Location/Qualifiers					
source	1..980					
	/organism="Corynebacterium glutamicum"					
	/db_xref="taxon:1718"					
	<1..960					
	/note="FRX00499"					
	/codon_start=1					
	/transl_table=11					
	/protein_id="CAC26116.1"					
	/db_xref="GI:12543465"					
	/translation="GNDILEKPTAEVVEYLRLADGIVRDAEALRKHLGVNOMNLG					
	OSFGFTTLYLHSLRADSLDNVITGSLAIDPADVANCYNNRRRSEEEYRFP					
	QREFFRGVLRNARAGEIVLPTGEVSETRLSIGHLGSDNGMDLYNLLEDPNS					
	AVYDLAGLLPGRNRPYLYVLIHSSYAGGVYTNMAAEVLPEDFEDPTLLTGGEVF					
	QEWTDVPSLTKPRKDVALLAQQEMPKLYDAKALENSQAKGAAYVANDVFVVDYSL					
	ETRAQHLPGVQLFTTSHENHGLRASSGAVLKLFLDLAHGREVR"					

BASE COUNT	215 a	267 c	264 g	234 t
ORIGIN				
Query Match	84.0%;	Score 16.8;	DB 9;	Length 980;
Best Local Similarity	90.0%;	Pred. No. 2.8e+02;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 caatgacggagaacgttcg 20
|||||
Db 307 CAATACGGGAACCTTCCG 326

RESULT 2						
AX065751	1404 bp	DNA				
DEFINITION	Sequence 95 from Patent WO0100842.	PAT	24-JAN-2001			
ACCESSION	AX065751					
VERSION	AX065751.1	GI:12543463				
KEYWORDS						
SOURCE						
ORGANISM	Corynebacterium glutamicum.					
	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;					
	Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;					
	Corynebacterium.					
REFERENCE	1 (bases 1 to 1404)					
AUTHORS	Pompeius, M., Kroeger, B., Schroeder, H., Zelder, O. and Habermayer, G.					
TITLE	Corynebacterium glutamicum genes encoding proteins involved in					
JOURNAL	homeostasis and adaptation					
	Patent: WO 0100842-A 95 04-JAN-2001;					
	Basf AKTIENGESellschaft (DE)					
FEATURES	Location/Qualifiers					
source	1..1404					
	/organism="Corynebacterium glutamicum"					
	/db_xref="taxon:1718"					
	101..1384					
	/note="RXN00499"					
	/codon_start=1					
	/transl_table=11					
	/protein_id="CAC26115.1"					
	/db_xref="GI:12543464"					
	/translation="MYGVSTPARNGSMFKTGLSGIQLEITLTLPEDVADEPTD					
	VFARIATRVGGEDLPYLVFLQSGPNEARPSLNPINPWLVALEEVVAVLDORGT					
	GRSTVAGNDILEKPTAEVVEYLRLADGIVRDAEALRKHLGVNOMNLGOSFGFTT					
	LHYSLRHADSLDNVITGSLAIDPADVANCYNNRRRSEEEYRFPRLQREFFRG					
	LVNARAGEIVLPTGEVSETRLSIGHLGSDNGMDLYNLLEDPNSAVYDLAG					
	ILPGRNRPYLYVLIHSSYAGGVYTNMAAEVLPEDFEDPTLLTGGEHVFQWMDTPV					
	SLKPRKDVALLAQQEMPKLYDAKALENSQAKGAAYVANDVFVVDYSLTRAQHLPG					
	VDLFTTSHENHGLRASSGAVLKLFLDLAHGREVR"					

CDS

BASE COUNT	302 a	384 c	384 g	334 t
ORIGIN				
Query Match	84.0%;	Score 16.8;	DB 9;	Length 1404;
Best Local Similarity	90.0%;	Pred. No. 2.8e+02;		
Matches 18; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy 1 caatgacggagaacgttcg 20
|||||
Db 731 CAATACGGGAACCTTCCG 750

RESULT 3

BACGIDA/c	BACGIDA	4161 bp	DNA	BCT	26-APR-1993
DEFINITION	Bacillus stearothermophilus	glycerol dehydrogenase (proposed gld)			
ACCESSION	gene, complete cds.				
VERSION	M65289				
KEYWORDS	glycerol dehydrogenase.				
SOURCE	Bacillus stearothermophilus (sub-species nondiastaticus) (library: lambda, EcoRI fragments) DNA.				
ORGANISM	Bacillus stearothermophilus				

Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
1 (bases 1 to 4161)
Mallinder, P. R., Pritchard, A. and Molt, A.
Cloning and characterization of a gene from Bacillus
stearothermophilus var. non-diastaticus encoding a glycerol
dehydrogenase
Gene 110, 9-16 (1992)
92184120

JOURNAL
MEDLINE
FEATURES
SOURCE

Location/Qualifiers
1. .4161
/organism="Bacillus stearothermophilus"
/sub_species="nondiastaticus"
/db_xref="taxon:1422"
/issue_1lb="lambda, EcoRI fragments"
/note="ORF1"
/codon_start=1
/transl_table=1
/protein_id="AA022476.1"
/db_xref="GI:551708"
/translation="MDILKITEKHREMERIKMECTFAEYLEILKEKPWIAOSASRV
YNMKDGIEIEINGRKRKFFNQEIFGLDELRLVEEYFPAAKRLDYKRRI"
742. .1854
/gene="gld"
742. .1854
/gene="gld"
/EC_number="1.1.1.6"
/codon_start=1
/transl_table=1
/product="glycerol dehydrogenase"
/protein_id="AA022477.1"
/db_xref="GI:142978"
/translation="MAAEVFTISPAKYVQGNVITIKIANLEGINKTVIADETVMK
IAGETIYNELKKNIAAEVYFSEASNEVERIANIKRKAALVIVGSGKITDTA
KAYDELDAIVITPAISTADAPTALSIVTISDGVFESYFVKNPDLVDTKITA
NAPRRLASGIADALATWEARSVIKSGGTAGGIPPIAAEAIAKCEQTLFKYGL
AYESKAKVVPALAEVAVEANTLLSGLFESGGAAAHAINGFALGELHILTHGE
KAVGTLVQALEESQOEIRYELIYSLDPTVLEDIKIKDKASREDILKAKAKATA
EGETIHNAFNTADVADAIFAADQYAAKAYEKHKR"
complement(2044. .2517)
/note="ORF3 is homologous to an ORF downstream of the spot
gene of E. coli; ORF3"
/codon_start=1
/transl_table=1
/protein_id="AA022478.1"
/db_xref="GI:142979"
/translation="MALHVLVPEPEIPANTGNIAATCATNTALHLIRPLGFTDDKM
LRRAGLDWEFVNHYDSDLDFQONONGDFYITKGRKYTSYDSDSKELFFI
FGRETTGLPKELAEENEDRCLRIPTENVRALNLSNTAAIIVYEAALROOSFHGLS"
complement(2550. .3455)
/note="ORF4"
/codon_start=1
/transl_table=1
/protein_id="AA022479.1"
/db_xref="GI:142980"
/translation="MDVOAYRYKINENKLOFLVNSVNNHROOSSELEWCEKKELAAK
RKAEIVATCLIKPKOIDEATVVDYLCORLLYKOGDPLXIEIGERAKESKNG
HLISDEVKLNKCFHKDEEDPMSENEHKKDRVSQYDRLKSVQAEAMNNSRNPNT
YFDVNCINYSQCLRAGDAPMRGYPNRSNGWLNANNNSYSVAHAILMLLSTSKTS
LRAEVSSPELLKPGVICYDPOGDGRDHSFTVAAKDSMPLVNNAOTSRRMYWS
YEDSTAYTENIOYKFFHIDNEDS"
complement(3511. .4161)
/partial
/note="ORF5: This ORF includes a region (aa23-103)
containing a potential iron sulphur centre homologous to a
region of Rhodospirillum rubrum and Chromatium vinosum;
putative"
/codon_start=1
/transl_table=1
/protein_id="AA022480.1"
/db_xref="GI:142981"
/translation="EFGSYVLGEMITNIPPPSPIEDRCGSCNICVDSPTGALVO

GGQDLSKICIAFLAQTGFLDEPRKIGNRLYGCOTQVCPENKKGDFHISEMER
DPEIAKPKLPLHLISNREFKFGHISGRKGPIDORMAIIALAHYKDTALPDLI
KLHEDSPRIIRGTAAWALIGRGDSVLELEKALEKRQDEVKNEIYKGIEMLKA"
BASE COUNT 1235 a 849 c 790 g 1287 t

Query Match 84.0%; Score 16.8; DB 2; Length 4161;
Best Local Similarity 90.0%; Pred. No. 2.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 caatgacgggaaacgtccg 20
Db 2141 CAATGACGGAAACGTTCCG 2122

RESULT 4
AF083905/c 674 bp DNA BCT 08-APR-1999
LOCUS
DEFINITION
Salmomella enterica VI strain RKS3057 major pilin protein fima
(fima) and fimbriin-like protein fimi (fimi) genes, partial cds.
AF083905
ACCESSION
AF083905.1 GI:4574180
KEYWORDS
SOURCE
ORGANISM
Salmomella enterica VI.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmomella.

REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE
1. .674
/organism="Salmomella enterica VI"
/strain="RKS3057"
/sub_species="VI"
/db_xref="taxon:59207"
<1. .486
/gene="fima"
/note="major pilin gene"
<1. .486
/gene="fima"
/codon_start=1
/transl_table=1
/product="major pilin protein fima"
/protein_id="AAD23933.1"
/db_xref="GI:4574181"
/translation="PTPVSVSGGTIHEFGKLVNACAVSTKSADQVTTLGQRTASFT
AIGNTTAQVPTIVLDCDDPKVAATAVAFGQADNTNNLLAASADNSATATGCVGI
EILNNTSSPLKPDGATFSAKOSIVEGTTLFTARYKATAAATTPGQANADATFIMKY
E"
561. .>674
/gene="fimi"
561. .>674
/gene="fimi"
/codon_start=1
/transl_table=1
/product="fimbriin-like protein fimi"
/protein_id="AAD23934.1"
/db_xref="GI:4574182"
/translation="MIRKGAALVGLITFPVIAQSVIYSGRVHLRGLYNG"

BASE COUNT 161 a 182 c 190 g 141 t

```
Query Match      80.0%; Score 16; DB 1; Length 674;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 caatcagcggaacct 16
        |||
Db       615 CAATGACGGGAACCT 600

RESULT 5
AF083912/c      674 bp      DNA      BCT      08-APR-1999
DEFINITION      Salmonella enterica VI strain RKS2995 major pilin protein Fima
                  (fima) and fimbria-like protein Fiml (fiml) genes, partial cds.
ACCESSION      AF083912
VERSION        AF083912.1 GI:4574201
KEYWORDS
SOURCE
ORGANISM
Salmonella enterica VI.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE      1 (bases 1 to 674)
AUTHORS        Boyd,E.F. and Hartl,D.L.
TITLE          Analysis of the type 1 pilin gene cluster fim in Salmonella: its
JOURNAL        J. Bacteriol. 181 (4), 1301-1308 (1999)
MEDLINE        99138752
REFERENCE      2 (bases 1 to 674)
AUTHORS        Boyd,E.F.
TITLE          Direct Submission
JOURNAL        Submitted (13-AUG-1998) Organismic & Evolutionary Biology, Harvard
                University, 16 Divinity Avenue, Cambridge, MA 02138, USA
FEATURES
source
1..674
/organism="Salmonella enterica VI"
/strain="RKS2995"
/sub_species="VR"
/db_xref="taxon:59207"
<1..486
/gene="fima"
/note="major pilin gene"
<1..486
/gene="fima"
/codon_start=1
/transl_table=1
/product="major pilin protein Fima"
/protein_id="AAD23947.1"
/db_xref="GI:4574202"
/translation="PTPVSYSGGTHIEGKLVNACAVSTKSADQVTLGQYRASF
AIGTTNQVPTIVLNDCKPRKVAATAVAATSGADNTNNLLAVSSADNSATATGVI
EILDNTSSPLKPGATFSAKQSLVEGTNLTFRATARKATTAATTPGOANADATFIMKY
E"
561..>674
/gene="fiml"
561..>674
/gene="fiml"
/codon_start=1
/transl_table=1
/product="fimbria-like protein Fiml"
/protein_id="AAD23948.1"
/db_xref="GI:4574203"
/translation="MIRSKALVGLITFPVIAOSVIVESGRVHLGKLVNG"
BASE COUNT      162 a 180 c 190 g 142 t
ORIGIN
```

```
Query Match      80.0%; Score 16; DB 1; Length 674;
Best Local Similarity 100.0%; Pred. No. 7.7e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 caatcagcggaacct 16
        |||
Db       615 CAATGACGGGAACCT 600
```

```
RESULT 6
AE000941/c      12099 bp      DNA      BCT      15-NOV-1997
DEFINITION      Methanobacterium thermoautotrophicum from bases 1724442 to 1736540
                  (section 147 of 148) of the complete genome.
ACCESSION      AE000941 AE0000666
VERSION        AE000941.1 GI:2623025
KEYWORDS
SOURCE
ORGANISM
Methanothermobacter thermoautotrophicus.
Methanothermobacter thermoautotrophicus.
Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
Methanothermobacter.
REFERENCE      1 (bases 1 to 12099)
AUTHORS        Smith,D.R., Doucette-Stamm,L.A., Delouhery,C., Lee,H., Dubois,J.,
                Aldredge,T., Bashirzadeh,R., Blakely,D., Cook,R., Gilbert,K.,
                Harrison,D., Hoang,L., Keagle,P., Lumm,W., Pothler,B., Qiu,D.,
                Spadafora,R., Ycaire,R., Wang,Y., Wierzbowski,J., Gibson,R.,
                Jiwani,N., Caruso,A., Bush,D., Reeve,J.N. et al.
                Complete genome sequence of Methanobacterium thermoautotrophicum
                deltaH: functional analysis and comparative genomics
J. Bacteriol. 179 (22), 7135-7155 (1997)
98037514
2 (bases 1 to 12099)
AUTHORS        Smith,D.R.
TITLE          Direct Submission
JOURNAL        Submitted (10-AUG-1997) Genomics and Technology Development, Genome
                Therapeutics Corporation, 100 Beaver Street, Waltham, MA
                02154-8448, USA
```

```
FEATURES
source
1..12099
Location/Qualifiers
/organism="Methanothermobacter thermoautotrophicus"
/strain="delta H"
/db_xref="taxon:145262"
/clone="MTH"
complement(64..960)
/gene="MTH1893"
complement(64..960)
/gene="MTH1893"
/note="Function Code:12.11 - Cell Processes,
Detoxification; similar to, pirLN:A64356 AC:A64356,
p(1)-1.8E-36, pId=198"
/codon_start=1
/transl_table=1
/product="cation efflux system protein (zinc/cadmium)"
/protein_id="AAB86353.1"
/db_xref="GI:2623026"
/translation="MESSEKIRLGRRAAFGIGCNVLLTSLNPLVIGSSGVALLVAEA
AHTLSVLTIVITYIGFRIGQRPDPKHPGHRALVGLVAVVEFLGIISTILSEA
YRKLFELEAPDDYATLMAFGIITANLMTYTRIGERINSPAIVAADQHOQVDEFS
EIAIMLGAVAGSHLGRFLDPLVAIVIAVLKTAFLVGRNVNIIAGAVSPIMEDI
EKASMSYDVKGGLHEVRINHEGPAVSVDLHIEVDGMLLDAIRIAHDERVIEDID
IKIVNVHVCPEKSKLEMS"
complement(962..2086)
/gene="MTH1894"
complement(962..2086)
/note="Function Code:5.01 - L-Amino Acid Metabolism,
Alanine--aspartate and glutamate metabolism; similar to,
sp.LIN:AA7.METHA AC:060317, p(1)-6.1E-82, pId=468"
/codon_start=1
/transl_table=1
/product="aspartate aminotransferase homolog"
/protein_id="AAB86354.1"
/db_xref="GI:2623027"
/translation="MIMKRTASRVKDTQSEIKRIFVADPTINIGIGPDSVPDH
VREAVKADVDEGLTHITSNKGMELEKRAIDKSKERVAEPESITVYGASEAIFM
CTGALLDIGHALIPDGLSYDACVRLSGAVSIIPVLSMDEGFSMPERVELITQD
TVRIINSPNPSTGSKDVKGIAEIAEDNDLIIISDEIKERIIYDGKHPAPQET
DNALVNGFSKTYAMTGLRIGYVAGCEDITEELKHYQVTAQASISQYALAAIAG
PNCVVKMVDPEFRRRDLMRSLTDMGLECVLPGLGAFYMPFPYAGDSSEFLKLSLEAGV
AVYPGSAFNGEGKIRIMSTAYTSTELIEAMEKIKYCVG"
```

```
gene complement(2121..3626)
/ gene="MTH1895"
CDS complement(2121..3626)
/ gene="MTH1895"
/ note="Function Code:14.01 - Unknown, Conserved protein;
similar to, gp:GI:1652365 LN:D90505, p(-)9.8E-41,
pid=35%"
/codon_start=1
/transl_table=1
/product="conserved protein"
/protein_id="AAB6355.1"
/db_xref="GI:2623028"
/translation="MRAVASCYSPSYSTAMSSLGHIITYHFLNSREDTWCERTIHPHR
SFESGSPLRDPEVVSFTIHYEDYFRIPAMFRCGLPDRDERGPLYIAGPCITSN
PLPSDFIDLEFIIGAEPEVLDMDRCLEDDPRDEFRIDRGVYDSPAERYIV
ENMEACHPVQIIVPTDDRSITPISGLSFGSLGSGRCRCRCMSGVLKPKRETS
LRKIDISESCFANTGYSKSLIGLAASDYSRIDELCNLMCLMAMSPSLRISVT
PELMALIRGLKKTITMAPESTMKLRDLINKPTDSWFEKTRKALEGLKMYKFLI
GAPESREDEEMARLMDLSEIRGRVSFVNPLIPKPHPLQWEGYDAREMQRIR
FLGRILVGLPYRMGSPRGILQHVLSGCGPEIGELIERASISGVPRWEMAHAGRPD
CARLPMHNISVGLREDFLMEBYIKMKDRLTPWCESGCGCMDSCLRH"
3797..4312
/ gene="MTH1896"
3797..4312
/ gene="MTH1896"
/ note="Function Code:14.01 - Unknown, Conserved protein;
similar to, pIR:LN:E64428 AC:E64428, p(-)3.3E-26, pid=36%"
/codon_start=1
/transl_table=1
/product="conserved protein"
/protein_id="AAB6356.1"
/db_xref="GI:2623029"
/translation="MFEKDDTKRRKYSLVAVGTFDFRHKGRLLDEAFVGETVMI
GVTSDERFAAGSEIGEPSCYVRKMLEYLRDKDDYHMRLLDDPYGTQDEAFPAIV
VSRETEPVAREINIRNRCPRELDITIDVNMDDGIPISSTRIRCEIDPMGHIIR
RIRGALRRRR"
4334..4939
/ gene="MTH1897"
4334..4939
/ gene="MTH1897"
/ note="Function Code:14.01 - Unknown, Conserved protein;
similar to, sp:LN:Y261_METUA AC:Q57709, p(-)1.8E-27,
pid=33%"
/codon_start=1
/transl_table=1
/product="conserved protein"
/protein_id="AAB6357.1"
/db_xref="GI:2623030"
/translation="MGCTFWYDDLAHIQVKEQDSVVEYMKVYVSGNPKVRATEN
VLGLIGNVEKGVESVGPDPQVGLSEETVRGAVNRARAFRCDELVSIGSLHRY
PETTGEVDLQWCAIYDGEHTTIGVSAGFEYPPVVEEVLAGEVGDVMDLTVGDEL
GRKGAVSFLSGGLDRGTNEGCVLMAIMPRMNPISLGLK"
5152..7170
/ gene="MTH1898"
5152..7170
/ gene="MTH1898"
/ note="Function Code:11.04 - Cell envelope, Murein
sacculus and peptidoglycan; similar to, pIR:LN:D64490
AC:D64490, p(-)2.6E-34, pid=31%"
/codon_start=1
/transl_table=1
/product="oligosaccharyl transferase"
/protein_id="AAB6358.1"
/db_xref="GI:2623031"
/translation="MROYMRFTLVAVIYVFGESIRIMNAGYDDPLTYDDGLOFQF
SDSTYNRRTENLAIRGIPGDVAVIGKIMDASHYPPGVPIDYPAIGILLAVILHAT
GYAVPSLRVCLIMPLAPLAVAGVICTFYRQGGDIPATIAELLVSAFVYLVKY
YGLVDTMLITLISMAVLEIFEADAREVNVPSAIGGFLIFLSTISWAGVYFY
TLVFEILITLTSGRKPLRNTIGNIPGHSVYPNFENLELIKILMPLHIDVY
DKSTVYMPDSTYVAELQPAYMKVGHISIPATVGAADIPAYMYKRRKSGSDP
VIFISWIATGALMLGCIKIRPMILPLPLVSACITFGILDCSRKEFRNKGLEFI
LIIFTVAQLAISGNMVAADLKPGYNDFEESAHIAETPDAVITTEMVSGHYVASE
ADRVVYVDGLAYIETLPARGVWYSSMDPEIPTTARDYMINLSSENRTLINILR
```

```
MLANSDDAYILKNQGSSEVQAVRINNELGLNESAMDYLMHRYGSESEALRILN
MTHPMTSRKIVVWVPEKMIIGLIASGSDGMDPAGYSLIRKDIVIEEELINSSGNT
VLEPTDGSYRINRKYRSLLELIMGVDENESIVYRNKQIKVYIINS"
7530..7703
/ gene="MTH1899"
7530..7703
/ gene="MTH1899"
/ note="Function Code:14.00 - Unknown, ; similar to,
sp:LN:YWLA_BACSU AC:P39150, p(-)0.98, pid=15%"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAB6359.1"
/db_xref="GI:2623032"
/translation="MYSAPCPASSKFNFLYTSINIHQFWLATLICVYPLINLADNT
LVILICSHQYHI"
7735..8286
/ gene="MTH1900"
7735..8286
/ gene="MTH1900"
/ note="Function Code:14.00 - Unknown, ; similar to,
sp:LN:IATP_PICJA AC:P09940, p(-)0.999, pid=06%"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAB6360.1"
/db_xref="GI:2623033"
/translation="MMPKRYNIDRYILEVLOGLDSRYELGERIRSEIFAVTDTIN
EAIKRLKASRIYTVGYDLYGVDCYVEROSIKPGCIYGLVORPVEKNLLIRLESE
NHHSESALNKLKRIKPMKTEGSELGVADAGISTYVNELISLDQDKRITITOKLVAALS
DEDDAPEQLRHLIYFEIRAGNF"
8626..10023
/ gene="MTH1901"
8626..10023
/ gene="MTH1901"
/ note="Function Code:14.00 - Unknown, ; similar to,
pIR:LN:S45091 AC:S45091, p(-)0.21, pid=06%"
/codon_start=1
/transl_table=1
/product="unknown"
/protein_id="AAB6361.1"
/db_xref="GI:2623034"
/translation="MIVLITITAGIWMIFACLYMSLGFATPPDIPMFLLFLFVPG
LIITLFLGKVDLKSLSKASEPKDKKRIKAGAPVAKEGGKQDRPGVSGIVYGE
KTEAPKPLKSTGKYQKRGCPVKVYRTSDEIKSRAIKPTVYTESDESIGAGESE
PSAASRIKPSSELSKEIIGELPEKPEKPSDETGAEAKPRKPAPIRPMAPAKKEV
PEEPAPPSREPEPKRFLRLPKIRARKEPTREPPAYKPAKKTTPAPAKPSRKA
DMSYTERLQRLKQDEYIENVDDVEDELDRLDSRKGANIRIATREPTSTWSBASD
VQDAMRETIITAAEERVVLMYPMIRINIDVSLKFKMDIESRLIODEASLDDASVELK
VLMNNVQIRTPMHIHTVAVAADKNGLIISTDPIYESFEVGIYKDKKSISEIKKLF
EEAMELSNEINLEGV"
10023..10715
/ gene="MTH1902"
10023..10715
/ gene="MTH1902"
/ note="Function Code:14.01 - Unknown, Conserved protein;
similar to, pIR:LN:B64445 AC:B64445, p(-)6.3E-48, pid=46%"
/codon_start=1

Query Match 80.0%; Score 16; DB 1; Length 12099;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatagcggaacgt 16
DB 10620 CAATGACGGGAACGT 10605

RESULT 7
LOCUS CELY65B4BR 4423 bp DNA INV 28-MAR-2000
DEFINITION Caenorhabditis elegans cosmid Y65B4BR, complete sequence.
ACCESSION AC024847
```

VERSION AC024847.1 GI:7140409
 KEYWORDS HTG.
 SOURCE Caenorhabditis elegans.
 ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 44237)
 AUTHORS The C. elegans Genome Sequencing Consortium, Washington University Genome Sequencing Center, St. Louis U.S.A. and the Sanger Centre, Hinxton, U.K.C.
 TITLE Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium
 JOURNAL Science 282 (5396), 2012-2018 (1998)
 MEDLINE 99069613
 REMARK Erratum: [[Published errata appear in Science 1999 Jan 1;283(5398):35 and 1999 Mar 26;283(5410):2103 and 1999 Sep 3;285(5433):1493]]
 REFERENCE 2 (bases 1 to 44237)
 AUTHORS Waterston, R.H.
 JOURNAL Submitted (01-MAR-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 44237)
 AUTHORS Waterston, R.
 JOURNAL Direct Submission
 TITLE Submitted (28-MAR-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT Submitted by:
 Genome Sequencing Center
 Department of Genetics, Washington University,
 St. Louis, MO 63110, USA, and
 Sanger Centre, Hinxton Hall
 Cambridge CB10 0RO, England
 e-mail: rwenematode.wustl.edu and jesusanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats: all regions were covered by sequence from more than one m13 subclone.

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder(P. Green and L. Hillier, ms in preparation).
 Location/Qualifiers
 1..44237
 /organism="Caenorhabditis elegans"
 /strain="Bristol N2"
 /db_xref="taxon:6239"
 /chromosome="I"
 /clone="Y65B4BR"
 1386..1387
 /note="5'UTR trans-splice site: see ESTs yk342b3.5, yk62a2.5 and yk516d1.5"
 1404..2089
 /gene="Y65B4BR.5"
 join(1404..1445,1495..1902,1952..2089)
 /gene="Y65B4BR.5"
 /note="contains similarity to alpha-NAC, non-muscle form proteins; coded for by C. elegans cDNA yk410q1.5; coded for by C. elegans cDNA yk361f9.5; coded for by C. elegans cDNA yk334b12.3; coded for by C. elegans cDNA yk655a1.3; coded for by C. elegans cDNA yk644h2.3; coded for by C.

misc_feature
 1386..1387
 /note="5'UTR trans-splice site: see ESTs yk342b3.5, yk62a2.5 and yk516d1.5"
 1404..2089
 /gene="Y65B4BR.5"
 join(1404..1445,1495..1902,1952..2089)
 /gene="Y65B4BR.5"
 /note="contains similarity to alpha-NAC, non-muscle form proteins; coded for by C. elegans cDNA yk410q1.5; coded for by C. elegans cDNA yk361f9.5; coded for by C. elegans cDNA yk334b12.3; coded for by C. elegans cDNA yk655a1.3; coded for by C. elegans cDNA yk644h2.3; coded for by C.

gene
 1404..2089
 /gene="Y65B4BR.5"
 join(1404..1445,1495..1902,1952..2089)
 /gene="Y65B4BR.5"
 /note="contains similarity to alpha-NAC, non-muscle form proteins; coded for by C. elegans cDNA yk410q1.5; coded for by C. elegans cDNA yk361f9.5; coded for by C. elegans cDNA yk334b12.3; coded for by C. elegans cDNA yk655a1.3; coded for by C. elegans cDNA yk644h2.3; coded for by C.

CDS
 1404..2089
 /gene="Y65B4BR.5"
 join(1404..1445,1495..1902,1952..2089)
 /gene="Y65B4BR.5"
 /note="contains similarity to alpha-NAC, non-muscle form proteins; coded for by C. elegans cDNA yk410q1.5; coded for by C. elegans cDNA yk361f9.5; coded for by C. elegans cDNA yk334b12.3; coded for by C. elegans cDNA yk655a1.3; coded for by C. elegans cDNA yk644h2.3; coded for by C.

elegans cDNA yk613d2.3; coded for by C. elegans cDNA yk678f10.3; coded for by C. elegans cDNA yk424f8.3; coded for by C. elegans cDNA yk584g3.3; coded for by C. elegans cDNA yk673h6.3; coded for by C. elegans cDNA yk117f10.3; coded for by C. elegans cDNA yk72e2.3; coded for by C. elegans cDNA yk484a9.3; coded for by C. elegans cDNA yk455d3.3; coded for by C. elegans cDNA yk483g7.3; coded for by C. elegans cDNA yk612d11.3; coded for by C. elegans cDNA yk617c3.3; coded for by C. elegans cDNA yk666e11.3; coded for by C. elegans cDNA yk345h7.3; coded for by C. elegans cDNA yk475d5.3; coded for by C. elegans cDNA yk623g12.3; coded for by C. elegans cDNA yk603f6.3; coded for by C. elegans cDNA yk62a2.3; coded for by C. elegans cDNA yk617c3.3; coded for by C. elegans cDNA yk603f6.3; coded for by C. elegans cDNA yk584g3.3; coded for by C. elegans cDNA yk428c12.5; coded for by C. elegans cDNA yk644h2.5; coded for by C. elegans cDNA yk267e1.5; coded for by C. elegans cDNA yk428c12.5; coded for by C. elegans cDNA yk628e1.5; coded for by C. elegans cDNA yk454d12.5; coded for by C. elegans cDNA yk517d10.5; coded for by C. elegans cDNA yk613d2.5; coded for by C. elegans cDNA yk473f10.5; coded for by C. elegans cDNA yk342b3.5; coded for by C. elegans cDNA yk72e2.5; coded for by C. elegans cDNA yk165h2.5; coded for by C. elegans cDNA yk655a1.5; coded for by C. elegans cDNA yk314d9.5; coded for by C. elegans cDNA yk345h7.5; coded for by C. elegans cDNA yk414a7.5; coded for by C. elegans cDNA yk117f10.5; coded for by C. elegans cDNA yk621e1.5; coded for by C. elegans cDNA yk678f10.5; coded for by C. elegans cDNA yk473f10.3; coded for by C. elegans cDNA yk62a2.5; coded for by C. elegans cDNA yk516d1.5"
 /codon_start=1
 /protein_id="AAE60854.1"
 /db_xref="GI:7332167"
 /translation="MTGSRFETROKKEPOVDSDSDNEAVROELTEORVAFAGSDHYILFGEAKIEDLTOHOMSAIENLKPTEAPDOKTVEEDNEVEDSTGERKDIELVISOANTTRKRAIRALKADNDIVNAIMSLTM"
 2734..7227
 /gene="Y65B4BR.4"
 join(2734..2839,2899..3156,3202..3431,4298..4507,5049..5323,5391..5604,5660..5580,6190..6258,6319..6579,6660..6881,6940..7227)
 /gene="Y65B4BR.4"
 /note="contains similarity to Pfam families PF00632 (HECT-domain (ubiquitin-transferase), score=403.2, E=2.5e-117, N=1) and PF00397 (WW domain, score=182.4, E=7.1e-51, N=4); coded for by C. elegans cDNA yk254h3.5; coded for by C. elegans cDNA yk339d4.5; coded for by C. elegans cDNA yk618c2.5; coded for by C. elegans cDNA yk52c9.5; coded for by C. elegans cDNA yk406b11.5; coded for by C. elegans cDNA yk48e8.5; coded for by C. elegans cDNA yk558g7.5; coded for by C. elegans cDNA yk503h4.5; coded for by C. elegans cDNA yk354f2.5; coded for by C. elegans cDNA yk361c11.5; coded for by C. elegans cDNA yk135a2.5; coded for by C. elegans cDNA yk191a2.5; coded for by C. elegans cDNA yk106b12.5; coded for by C. elegans cDNA yk106b12.3; coded for by C. elegans cDNA yk616g10.5"
 /codon_start=1
 /protein_id="AAE60858.1"
 /db_xref="GI:7332171"
 /translation="MARNEPSSQPPSSGSGNGTPAOQNGSARPKVTVKVNAPFKAADCVEITSDPTSSAARKTKTVKKTAPAKENNELNHAHNESSISFRLLOKALFEDTCLGMALIKLSILFRNENGEKFNINNSILAKRSLITGNTINIFSGYPERKRSACVAFETASASSPSATSGVATSSSARAPATKROLTAPSTAAANAAATGPPAACAAEOLPDGEMKRFDOYGRKYVYDHTKSTTMRPSTOPLPGCEMRPRRGVRYVDINTRTTWOPKPTADMLEAHEOWSGDQAMLDOEFLIQNNFSADDDPLGIPBEWEKQDPTTSRMYFVNHNARTTOWEDPRTQGRFSDOPLPGMEWREFTDQEPFTIQSXTTYNDPRTGRVGLGVGVOMAEKSFPMKIAQARFLCISNPNVAKITVSNNVFE DSFOELMRKAVDLRRRLVIOERGEIDGCAVARMFLLSHEVLPNPMVCLPMYNG NNSLIQINPASFVNPHIKYEEIGRIFMIRKNGMDLKPESDOYISPSPKNALPHG KFIYSGTFMRFFKMLNKIKVLKLDIOGVSEIINSLMKNKDNIDCDMELTVADYE LIGELTYELKEGTEIAVTEENKLEYIELVEMRNPNGVEOTKAFFTGFSNVFPLE

gene
CDS
16304..21657
/gene="Y65B4BR.3"
join(16304..16508,17309..17505,18001..18414,18731..19181,
19416..19990,20620..20693,20741..20927,20978..21361,
21457..21657)
/note="Y65B4BR.3"
/gene="contains similarity of membrane protein
PATCHED-related proteins"

gene
CDS
26298..26394,26997..27159
/gene="Y65B4BR.2"
join(21953..22235,24000..24139,25103..25259,25302..25463,
26298..26394,26997..27159)
/gene="Y65B4BR.2"
/codon_start=1
/evidence="not_experimental
/product="Hypothetical protein Y65B4BR.2"
/protein_id="AAF60856.1"
/db_xref="GI:7332169"

gene
CDS
29103..29176,30205..30706
/gene="Y65B4BR.1"
join(28359..28465,28516..28715,28760..28926,28970..29053,
29103..29176,30205..30706)
/translation="MENSATSOFTHQKLLAPAFHDLFFPPPOFLFEISEFPOL
VECHRLVLFSSAGSPPEPTDLRLVPLIFSSILVTOIVPSLDVLSIFSPAGVRF
FTLPVGRNVPIRLFEVNAACRPKSKPNVDLANLQMDVATKAYSLF
FSGMDQIEKFMKAVYVDSKEVHKEDCSIEYEDMLOTPTATFTSKOYGVINDV
VTNMGDSMGVPEPGAVLITGSESDGCPPEPRFVGLSDGEQOYMLISPLKRPIM
VLTRDLTLFEETKMKREYVDPEKNGFSPMALNRLHFTVDVCRKYNKLENGNV"
28359..30706
/gene="Y65B4BR.1"
join(28359..28465,28516..28715,28760..28926,28970..29053,
29103..29176,30205..30706)

Query Match
Best Local Similarity 80.0%; Score 16; DB 6; Length 44237;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 aatgacgggaacgtt 17
|||||
Db 16259 AATGACGGGAACGTT 16244

RESULT 8
AC006889/c
LOCUS
DEFINITION
AC006889 267118 bp DNA HTG 26-FEB-1999
6 unordered pieces.
AC006889
VERSION
AC006889.2 GI:4309908
KEYWORDS
HTG: HTGS, PHASE1.
ORGANISM
Caenorhabditis elegans.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
1 (bases 1 to 267118)
Waterston, R.H.
The sequence of *Caenorhabditis elegans* clone
Unpublished
2 (bases 1 to 267118)
Waterston, R.H.
Direct Submission
Submitted (24-FEB-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA

COMMENT

On Mar 1, 1999 this sequence version replaced gi:4263459.
* NOTE: This is a "working draft" sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
source
1
1
5876
5889: contig of 5875 bp in length
5890
13618: contig of unknown length
13619
13632: gap of 7729 bp in length
13633
28840: gap of unknown length
28841
28854: gap of 15208 bp in length
28855
110971: contig of 82117 bp in length
110972
110985: gap of unknown length
110986
174804: contig of 63819 bp in length
174805
174818: gap of unknown length
174819
267118: contig of 92300 bp in length.
Location/Qualifiers
1..267118
/organism="Caenorhabditis elegans"
/db_xref="taxon:6239"
/clone="Y65B4"

BASE COUNT
83652 a 50557 c 48664 g 84175 t 70 others
ORIGIN

Query Match
Best Local Similarity 80.0%; Score 16; DB 60; Length 267118;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 aatgacgggaacgtt 17
|||||
Db 128243 AATGACGGGAACGTT 128228

RESULT 9
AE005077
LOCUS
DEFINITION
AE005077 AE004437
VERSION
AE005077.1 GI:10581157
KEYWORDS
SOURCE
ORGANISM
Halobacterium sp. NRC-1.
Halobacterium sp. NRC-1
Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
Halobacterium.

REFERENCE

1 (bases 1 to 14625)
Ng, M.V., Kennedy, S.P., Mahairas, G.G., Bergquist, B., Pan, M.,
Shukla, H.D., Lasky, S.R., Baliga, N., Thorsson, V., Shrogha, J.,
Swartzell, S., Weir, D., Hall, J., Dahl, T.A., Weltl, R., Goo, Y.A.,
Leitner, B., Keller, K., Cruz, R., Danson, M.J., Hough, D.W.,
Maddocks, D.G., Jablonski, P.E., Krebs, M.P., Angevine, C.M., Dale, H.,
Isenbarger, T.A., Peck, R.F., Pohlisch, M., Spudis, J.L.,
Jung, K.-H., Alam, M., Freitas, T., Hou, S., Daniels, C.J., Dennis, P.P.,
Omer, A.D., Edhardt, H., Lowe, T.M., Liang, P., Riley, M., Hood, L. and
DasSarma, S.

AUTHORS

TITLE
JOURNAL
PUBMED
REFERENCE
From the cover: genome sequence of halobacterium species NRC-1
Proc. Natl. Acad. Sci. USA 97 (22), 12176-12181 (2000)
11016950
2 (bases 1 to 14625)

AUTHORS Ng,W.V., Kennedy,S.P., Mahairas,G.G., Bergquist,B., Pan,M., Shukla,H.D., Lasky,S.R., Baliga,N., Thorsson,V., Sbrogna,J., Swartzell,S., Weir,D., Hall,J., Dahl,T.A., Meli,R., Goo,Y.A., Leithauser,B., Keller,K., Cruz,R., Danson,M.J., Hough,D.W., Meddocks,D.G., Jablonski,P.E., Krebs,M.P., Angevine,C.M., Dale,H., Ienbarger,T.A., Peck,R.F., Pohnschod,M., Spudich,J.L., Jung,K.H., Alam,M., Freitas,T., Hou,S., Daniels,C.J., Dennis,P.P., Omer,A.D., Ehardt,H., Lowe,T.M., Liang,P., Riley,M., Hood,L. and Dassarma,S.

Direct Submission
Submitted (14-JUL-2000) Institute for Systems Biology, 4225 Roosevelt Way NE, Seattle, WA 98105, USA

FEATURES
Location/Qualifiers
1. .14625
/organism="Halobacterium sp. NRC-1"
/strain="NRC-1"
/db_xref="taxon:64091"
complement(83. .153)
/gene="trn29"
/product="tRNA-Gly"
complement(83. .153)
/gene="trn29"
401. .1252
/gene="VNG1688C"
401. .1252
/gene="VNG1688C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=11
/product="Vng1688c"
/protein_id="AAG19935.1"
/db_xref="GI:10581158"
/translation="MTRSVLVPSLVLEADKKREARKLGVARAARFIDRVVPP
DEDEGRMGGEFETVLRATPPYLKREAFDRDELAVGLPLRLSSMTGSDSSG
SGSLRQIVTVGSEGRVNCQMOHPLSHPPGMASEGERTYIRVSRPVRAL
VDDPLGEFSVERTGIDALDRSAGVRIATSRHGCEPLSVASLGGYRRIARDCVTAF
GABERCLPPMLGVSADEVNESVTDSSADAPARDAMLNTIPDQSEVVRTEAVLATL
GSJFLYE"
1257. .2264
/gene="trp13p"
/note="VNG1689G"
1257. .2264
/gene="trp13p"
/note="Rpl3p"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L13p"
/protein_id="AAG19936.1"
/db_xref="GI:10581159"
/translation="MPQPNRPKSGSGFSFRKRAESVPRFNSMPADGVEVGLQGFAG
YKAGMTHVLVDKNAKAPBGMETVAVTVETPPMRAAARVLEDPYQKPLTEVM
ADPTHESIDLTGVDPDGDTDELIEALDTEELADIVITHVPGDAGVKKPLTEVM
ETRVGGGTLADRLFEFADLIEDGVHAFDGVFRAGFTDAAGITKGGTQGVPRMIV
QKRGKARQGRMRIGLNGMPNSRVRSYVPOOGCGVQHORTEINKRLDINDGDP
TPDGGPNTGCEVDPITLVKGSYVGPEDQLVRRPRANRPNESRLDPEVRYVSTASNO
G"
2268. .3020
/gene="rpl4e"
/note="VNG1690G"
2268. .3020
/gene="rpl4e"
/note="Rpl4e"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L4E"
/protein_id="AAG19937.1"
/db_xref="GI:10581160"
/translation="MOTVVRDLDDAGTLDLPRVVEEPPRDLVKRAVLAQAARNTQ
EYGADEAGLRITAEQSGSGMAHYKANGOGARVPQTVGGRKAHPKAEKHGLDY
NDEKRAAYRAAAATTDSELVADRGHNFDDDEPPLVVSDDFEDLVKTKDDVSLLEA
LGVHADIERADEGRIVRAGCGTLRGKYOCEPTSLITVTAESGSPSTAAKRLAGDVAT
GREVNADDLAPGAEPPGRITVWTSAAVEEVAQR"
3017. .3271

CDS
/gene="rpl23p"
/note="VNG1691G"
3017. .3271
/gene="rpl23p"
/note="Rpl23p"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L23p"
/protein_id="AAG19938.1"
/db_xref="GI:10581161"
/translation="MSGIIDYPLVTEKAMDEMPQNKQPIYVIDAKPEIRDVSESE
IDVTVVDVNTQITPEAKKATVKLASDDQDVASRLGVF"
3274. .3996
/gene="rpl2p"
/note="VNG1692G"
3274. .3996
/gene="rpl2p"
/note="Rpl2p"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L2p"
/protein_id="AAG19939.1"
/db_xref="GI:10581162"
/translation="MGRRIQGRGRGRTSTFRAPSHRYKAELSHKRTEDTVLAGEVI
DVEHDPARSAFVARVAFEDDQRLVLASEGVGDTLEIGISATIEBGNLPLAEIPE
GVPCVNESSHPRGDKGFARAGVNADLVYTERATIVLEPSGETKRLSPDCRATIGV
AGGRTKEKPFVKAGNKKMKKARGTKWPRVGVAMNAVDPFGGGGRHHPGRKSVSR
DAAPGRKVDIASKRTGCGNE"
4000. .4422
/gene="rps19p"
/note="VNG1693G"
4000. .4422
/gene="rps19p"
/note="Rps19p"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S19p"
/protein_id="AAG19940.1"
/db_xref="GI:10581163"
/translation="MSPEYRTGRGREGFTYRGHDLDELQEMSLBDVAELLPARQRTIT
RGLSEHHKVLAEAREGTEETANNPTRTHLRMPVLPFVGLTPAVYTTQGEEREREV
QPMIGHYLGEPQLTRSSVREHQAQIGATSSKFPVPLK"
4426. .4896
/gene="rpl22p"
/note="VNG1695G"
4426. .4896
/gene="rpl22p"
/note="Rpl22p"
/codon_start=1
/transl_table=11
/product="50S ribosomal protein L22p"
/protein_id="AAG19941.1"
/db_xref="GI:10581164"
/translation="MGISYVDVDSASAKMLRERSISLKHSAIAREISGFTVADA
KEYLQAVTIDERSVYPRKQHSYGVGRHNDIDGMARGRIPKASADFLKLSNVENPADQ
QGFDADEMTVEIHAAPKRVSGSQRKPRAMGRATVWNTLCTDVEIVTEEVTA"
4896. .5804
/gene="rps3p"
/note="VNG1697G"
4896. .5804
/gene="rps3p"
/note="Rps3p"
/codon_start=1
/transl_table=11
/product="30S ribosomal protein S3p"
/protein_id="AAG19942.1"
/db_xref="GI:10581165"
/translation="MADELEFIEQGLQRSQIDEFFAEELARAGYGGNELAPTPMGNOI
VLAEKRPAGVIGIKGKIRKITTOLEBRFPLEDPDQVDOVEEPDLNAQIVARLNA
LERGAPFRKAGHTTIRIMSGALGAETVLSGVTRGQGVKAFNGGYIKHNEPPAE
IVHGGVAMKIGTIGVNYKTIIPNAELPDPEIQEDADIEDLVYDEARAGDDEEL
LEGEDADADADADAAREPSEPADEFEEVETDDEVEBELDELADAVEGEDDEDEF
SDVDEADAPYAEDDLDEMDDEDGAE"

gene 5804. .6019
/gene="rp129p"
/note="VNG1698G"
5804. .6019
/gene="rp129p"
/note="Rp129p"
/codon_start=1
/transl_table=1
/product="50S ribosomal protein L29p"
/protein_id="AAG19943.1"
/db_xref="GI:10581166"
/translation="MAILVYSEIRDMTPAEREAELELRTELNSKAVKAAGAPDNP
GRISLKKTIARIKTVOREGDLADDE"
6028. .6339
/gene="VNG1699C"
6028. .6339
/gene="VNG1699C"
/note="conserved hypothetical protein"
/codon_start=1
/transl_table=1
/product="Vng1699C"
/protein_id="AAG19944.1"
/db_xref="GI:10581167"
/translation="MTTTPESLPRLHGLHARVADTDGSRVGIAGRVVDETMTQV
VLRASGVAVQPKKATFEPLTDEAAPNGVGTAKFPAGETROTGESAVVTV
GGRLSRPERSENGVDSKMR"
6390. .6719
/gene="rps17p"
/note="VNG1700G"
6390. .6719
/gene="rps17p"

Query Match 79.0% Score 15.8; DB 1; Length 14625;
Best Local Similarity 89.5% Pred. NO. 8.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 caatgacgggaacgctcc 19
||||| ||||| ||||| |||
Db 8856 CAATGACGCCAACGATCC 8874

RESULT 10
LOCUS MTV006 22440 bp DNA BCT 18-JUN-1998
DEFINITION Mycobacterium tuberculosis H37Rv complete genome: segment 54/162.
ACCESSION AL021006 AL123456
VERSION AL021006.1 GI:3242277
KEYWORDS
SOURCE Mycobacterium tuberculosis.
ORGANISM Mycobacterium tuberculosis
Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
Actinomycetales; Corynebacteriaceae; Mycobacteriaceae;
Mycobacterium.
1 (bases 1 to 22440)
Cole,S.T., Brosch,R., Parkhill,J., Garnier,T., Churcher,C.,
Harris,D., Gordon,S.V., Eiglmeier,K., Gas,S., Barry III,C.E.,
Tekala,F., Badcock,K., Basham,D., Brown,D., Chillingworth,T.,
Connor,R., Davies,R., Devlin,K., Feltwell,T., Gentles,S.,
Hamlin,N., Holroyd,S., Hornsby,T., Jagels,K., Krogh,A., McLean,J.,
Moule,S., Murphy,L., Oliver,S., Osborne,J., Quail,M.A.,
Rajandream,M.A., Rogers,J., Rutter,S., Seeger,K., Skelton,S.,
Squares,S., Squares,R., Sultison,J.E., Taylor,K., Whitehead,S. and
Barrell,B.G.

TITLE Deciphering the biology of Mycobacterium tuberculosis from the
complete genome sequence
JOURNAL Nature 393 (6685), 537-544 (1998)
MEDLINE 98295987
REMARK Erratum: [[Published erratum appears in Nature 1998 Nov
12:396(6707):1901]
2 (bases 1 to 22440)
AUTHORS Parkhill,J.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-1998) Submitted on behalf of the Mycobacterium

COMMENT
tuberculosis sequencing and mapping teams, Sanger Centre, Wellcome
Trust Genome Campus, Hinxton, Cambridge CB10 1SA unite de Genetique
Moleculaire Bacterienne, Institut Pasteur, 28 rue du Docteur Roux,
75724 Paris Cedex 15, France E-mail: parkhill@sanger.ac.uk
On Jun 20, 1998 this sequence version replaced gi:2695814.
Notes:
Details of M. tuberculosis sequencing at the Sanger Centre are
available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/M_tuberculosis/) CDS have
been renumbered from the original cosmid submissions but the old
gene designations are in brackets after the new gene numbers.
Gene prediction in Tbpase (Krogh) supplemented with visual inspection
of positional base preference in codons, especially where there is
an increase in the observed/expected third position G + C.
CAUTION: In some cases we may not have predicted the correct
initiation codon. Where possible we choose an initiation codon
(atg, gtg, or ttg) which is preceded by an upstream ribosome
binding site sequence (optimally 5-13bp before the initiation
codon). If this cannot be identified we choose the most upstream
initiation codon.

FEATURES
source
location/Qualifiers
1. .22440
/organism="Mycobacterium tuberculosis"
/strain="H37Rv"
/db_xref="taxon:1773"
<1. .>22440
/note="fragment designated v006. Does not represent a
physical clone"
108. .112
/note="possible RBS, agagg, for RV1234"
124. .651
/gene="RV1234"
124. .651
/note="RV1234"
/score=0.929"
/codon_start=1
/transl_table=1
/product="hypothetical protein RV1234"
/protein_id="CAI15890.1"
/db_xref="GI:2695820"
/db_xref="SPTREMBL:O50451"
/translation="MTSPFQRPVPGSTPAAGAGRGVPAIPPPKGVSPYTA
EAORAVDLSQDFPVQOVTIVGDLMQVEVTRTLMPKVLGGVSLGALGIFGL
VLGFSPMPMSALVTGLVAGVFPLITISAVPYNARGRDPSSTMQLVAGRYDVLCP
ONAKARDLARLAI"
663. .666
/note="possible RBS, gaggg, for RV1235"
672. .2078
/gene="1pgy"
672. .2078
/note="1pgy"
/gene="RV1235, (MTV006.07), len: 468. Unknown lipoprotein
highly similar to M. leprae protein u1756v (469 aa),
possibly involved in sugar (e.g. maltose) uptake. FASTA
scores: MU01518034opt: 2442 E(): 0: (77.4% identity in 470
aa overlap). Similar to P18815(MAL:ENTAE MALTOSE-BINDING
PERIPLASMICPROT (396 aa)FASTA scores: opt: 193 E():
2.3e-05; (24.2% identity in 297 aa overlap). Contains
PS00013 Prokaryoticmembrane1lipoprotein lipid attachment
site. Tbpase score 150.914"
/codon_start=1
/transl_table=1
/product="1pgy"
/protein_id="CAI15891.1"
/db_xref="GI:2695821"
/db_xref="SPTREMBL:O31095"
/translation="MWSRGRIPRLGAVALVLTAAAGCAGSGGLVSEFYTPATG
ATPFAIORCNOGRCGRFTTAOVSLPSPNOROLRLTGNDRDTQWALVWTA
EFAAGKALPLISDDPAIENDAVADLPPLATKAGNNHKLITAAPYTNTOLAKYRD
LVNSPTDWMAMIAEARLHAAGPSPVIAVANGDEGLVWVFNLVASAGSVLSEDG
RHVNLTDTPAHRAVTSALQILKSVATTPGADPSITREBGSRLAEDEGKAALVWVW

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 8, 1998 this sequence version replaced gi:2291178.
Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rwenematode.wustl.edu and jesesanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded or sequenced with an alternate
chemistry; an attempt was made to resolve all sequencing problems,
such as compressions and repeats; all regions were covered by
sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is C53A3, 700 bp overlap; 3' cosmid is ZC250, 200 bp
overlap. Actual start of this cosmid is at base position 501 of
CELF44E7; actual end is at 16254 of CELC250

NOTES:

Coding sequences below are predicted from computer analysis, using
the program GenefinderP. Green and L. Hillier, ms in preparation).

FEATURES
SOURCE

Location/Qualifiers

1..22480

/organism="Caenorhabditis elegans"

/strain="Bristol N2"

/db_xref="taxon:6239"

/clone="F44E7"

/chromosome="V"

255..1363

/gene="F44E7.2"

/join(255..417,468..1140,1192..1363)

/gene="F44E7.2"

/note="Similar to nltrophenyolphosphatase: coded for by C.
elegans cDNA yk312h2.5; coded for by C. elegans cDNA
yk267e8.5; coded for by C. elegans cDNA yk312h2.3"

/codon_start=1

/protein_id="AAC25793.1"

/db_xref="GI:3294491"

/translation="MKYPTITDSOOLPREPHDPLKAGLDPKCRSTKKPLCDPTPAKVK
TIDFTIDADGVLMGESSVMPGSPRLIDYLVKHKQITVLTNNKTSRAVYAKKIAVK
GYNSKMNKNLVPAVAVADTLHRAIGDGRVLLIGOGRLDEMDLGIEYFGGPE
KKQEDAGSGAFMYDILKLEENYGAVVYGEYFKYIKMKASNYLRKGVFLVATNED
ETGCGPNEVYIPAGPIVAIKCAGSDPLTVKPCPAVNYIKRKININPSTPMI
GDRNTVYKFGDRHGMKTLVLVSGCHQIEDIIEQNMRDMVDPYVAPCGALVPER
I"

2001..2253

/gene="F44E7.3"

/join(2001..2045,2095..2253)

/gene="F44E7.3"

/codon_start=1

/evidence="not_experimental"

/protein_id="AAC25788.1"

/db_xref="GI:2291180"

/translation="MTLQLOELMRQSDKEIEVLQSRLELAQAASDLTDRILHEDLRA
SDARQETFKFLPQPRNOKME"

complement(2795..6738)

/gene="F44E7.4"

complement(join(2795..3047,3095..3431,3478..4078,
4348..4482,4530..5338,5385..5632,5685..6253,6307..6451,
6632..6738))

gene
CDS

gene
CDS

/gene="F44E7.4"

/note="strong similarity to peptidase family M16 or the
insulinase family: coded for by C. elegans cDNA yk18e4.3;
coded for by C. elegans cDNA yk6c6.3; coded for by C.
elegans cDNA yk87d9.3; coded for by C. elegans cDNA
yk133a3.3; coded for by C. elegans cDNA yk34h4.3; coded
for by C. elegans cDNA yk15f11.3; coded for by C. elegans
cDNA yk5a5.5; coded for by C. elegans cDNA yk182f7.5;
coded for by C. elegans cDNA yk87d9.5; coded for by C.
elegans cDNA yk133a3.5; coded for by C. elegans cDNA
yk18e4.5; coded for by C. elegans cDNA yk15e2.5; coded
for by C. elegans cDNA yk17b6.5; coded for by C. elegans
cDNA yk6e1.5; coded for by C. elegans cDNA yk179h10.5;
coded for by C. elegans cDNA yk34h4.5; coded for by C.
elegans cDNA yk6c6.5"

/codon_start=1

/protein_id="AAC25789.1"

/db_xref="GI:2291181"

/translation="MSLSLSEKFGVSPALAHAKRTPESSVNFSQLRRPNSSLR
LRVYNINSNPLPKMTAGKNIVLRHDLVKGADAREYRGLETIGIVLVSDP
TTDKSAALDVKGVLMDPWLPGIAHCEHMLFGTAKYSEDESKFLAAGSSN
AYTSSDHYHEDVKKPDLPSALDRFVQFSLSPOTETRETCADVSESNLND
LMRFQVDRSRKPEGHDIGKFGTGKQTLLEDARKKIEPPDALLQFKKAYSSDIAT
CCIVGKEPLNVEISLTLEPDALENKVERKVEEPEYGPDLAKRIDVVPIDTRL
VSISEPPDLNGEFLSOPGHYISHLIGHEGCSLSLELRGLGWSLSQSDHTQAG
GVYNTMDLSTEGEIVDEIITLMFNVIGMOSAGPKQWVDELAESAVERKDK
OPMTAIVASISLOIPEPHILSSRYLTITKYEPRIRKLLSMLSPANVQVRSOKK
GQEBNTNPPVYGTETKAYTDISPEPKKAKENLAKTSHALHLPENYEATNPDDKPRE
SVKNHPRILISDDGSRVWERKODDETNPQKQETLALTPVAVQNPRLNSLMLWC
LSDLTAEETVADLAGLCKQLSESPFQVQKVSNNREVAHSLTLHAYGDEQALF
AKHLANRMTNFKIDKTRDVLFEESLKRALTNHAFSQPYLQHNQULIVKWSAKQ
LLAVCDSTLEDOVGFAKEMQAEFMELFVHNSSEKALQSKELMDLVKSAVNSR
PLVREHNPRELONNGDEVYVRLQTHVGCVEYVQIGONTYDNAVAVGLIDOL
IREPAFNTRPNEALGYIVMGVSRNGCTVALIYOGPKSVHLEIEVFLSVRK
ELAEPRQEPFNQVSGMIAERLEKPKTILSSFRFPNMBIEROVNFARREEVALLIK
IKDDVLELFDKRIKDAERKRAVPHGNEDEAVNTIKNAESGKKEKYLVS
DQLRQFLPLVGRVADLKPIDGVDPLEHQTTSKY"

complement(8821..10572)

/gene="F44E7.5"

complement(join(8821..9324,9577..9913,10129..10244,
10441..10572))

/gene="F44E7.5"

/codon_start=1

/evidence="not_experimental"

/protein_id="AAC25790.1"

/db_xref="GI:2291182"

/translation="HELSTDLMEKISYSLADHPSPCLLSSSSKRLRELNOFVIGHVE
ALHLITLPKDECCDEKSKSVQPEMLLNKSGSRKRAHFHSIFCLDIDILAVN
DLLEFAKELITLRFNNPENSVDFTVFNKSLKRCEDICFNGSLPVHNGVSDMFHSC
HNMRKLVISNPTDIFSQFWPKLGSVEFLDLMIHSEDELAINSKLYKRLSVYE
FCFSKATITIDYATIAFLOKCDQVCKCFNFGHSCVAPVIVARLLDIPRSIVFGCVA
CRSGETVYKGGFMENGVNAVVOYVTKILLERDFSLVGTILADAVLPERPNANSVS
MYIQKDSVCGPNCIDIHIEISQKAGVL"

complement(11113..12353)

/gene="F44E7.6"

complement(join(11113..11319,11364..11756,11874..11997,
12043..12129,12177..12355))

/gene="F44E7.6"

/codon_start=1

/evidence="not_experimental"

/protein_id="AAC25791.1"

/db_xref="GI:2291183"

/translation="MDPRALRIVCFLOHLOGIEFVNAYKAMSKYFGDHMEYRHFEYV
MYGLANGSLTVDSYDRPHTSERKEFTGTVOVAMILDKLPIDLFARAACRPDKI
IDNOKSTKRTIMNMLDERVRLTNNISIRGLINDCPENLVMMWMDDEVYRIEFLAIKGY
LESTIGLSAKSVLNMDDTYETIAELQHVAKREYLEDLTIKLGSERELVGLDEMKKA
KNLKLAKVPHSITIVKHLNHRKSNMFDVSGFSRDDIIRKREDEVLTKSTFETSE
NWKFSARLVAVFVFDHSTPSSSGSIVYKSKNGTPIADVHPNGFTIKRKIDSNH"

complement(12689..15444)

/gene="F44E7.7"

complement(join(12689..12910,12966..13058,13111..13179,
13726..14592,14894..15112,15159..15222,15357..15444))

/gene="F44E7.7"

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

gene
CDS

/note="weak similarity to Na+-dependent inorganic phosphate cotransporters"
/codon_start=1
/evidence=not_experimental
/protein_id="AAC25792.1"
/db_xref="GI:2291184"
/translation="MNNNSKILFHHIFDISITSLASTSKRTNVRHISSEVFHFKTR
YLIMILSLICMTVGOMNSUSFNFTVICMEDVVDVDFSNRNTDLHWLEDPQSKMIFSG
VALGAVIGLLPLPMDNIGLITFTFLGISASSLLEPLSVHMGFYAVFIVRILOG
IGTSILYTVAVYIPGIMAPKTEMTGTLAVSCGFOLSNIIICMPVSGILCSDGMMPDI
YIFGSLTILYVVFELFVADAPKRNHVSCKSMICAEKKRCKKESVPRALICED
KYLATWLSMCGRNAFAYVLYXGPLYRVLHEDYKGTGCMALPFGSAIYKFGSG
QLDRLTMSSEKTRFYTCILISIGLAAGVWSLTSRTIAOISTFTFANTLGCTITM
GVVCKQLRCQOHVHFAVYVAVAFMACIOWFLVGLGVILCPNPTPEWSPFLFIVSGI
VLIVNIPFPEFLTAAADYAKRP"
/gene="F44E7.9"

Query Match 79.08; Score 15.8; DB 5; Length 22480;
Best Local Similarity 89.5%; Pred. No. 8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 caatgacgggaacgttc 19
|||||||
Db 21626 CAATGACGGAAGACGTTCC 21608

RESULT 12
LOCUS CELC06G1 31205 bp DNA INV 30-NOV-1995
DEFINITION Caenorhabditis elegans cosmid C06G1.
ACCESSION U41014
VERSION U41014.1 GI:1086637
KEYWORDS
SOURCE .
ORGANISM Caenorhabditis elegans strain-Bristol N2.
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 31205)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, M., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
Fullon, L., Gardner, A., Green, P., Hawkins, J., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,
Saunders, D., Showkeen, R., Smaildon, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstock, L.,
Wilkinson-Sprat, J., and Wohldman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans

JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE 2 (bases 1 to 31205)
AUTHORS Wu, X and Le, T.
TITLE The sequence of C. elegans cosmid C06G1
JOURNAL Unpublished (1995)
REFERENCE 3 (bases 1 to 31205)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1995) Robert Waterston
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1HQ, England
e-mail: twenematode.wustl.edu and jes@sanger.ac.uk
NEIGHBORING COSMID INFORMATION:
The 5' cosmid is F59C12, 1000 bp overlap; 3' cosmid is F43B10.
Actual start of this cosmid is at base position 1 of CELC06G1.

NOTES:
Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in preparation).
location/Qualifiers
1. 31205

FEATURES
source
/organism="Caenorhabditis elegans"
/db_xref="taxon:6233"
/map="X"
4029. .6783
/gene="C06G1.4"
join(4029. .4209,4253. .4405,4463. .5281,5407. .5498,
5543. .6058,6619. .6783)
/gene="C06G1.4"

CDS
/note="coded for by C. elegans cDNA yk67e8.3; coded for by
C. elegans cDNA yk84g7.3; coded for by C. elegans cDNA
yk84g7.5; coded for by C. elegans cDNA yk67e8.5; coded for
by C. elegans cDNA yk61b1.5; Similar to D. melanogaster
homeotic gene regulator (Brahma protein) (weak)."
/codon_start=1
/protein_id="AA82305.1"
/db_xref="GI:1086638"

/translation="MESLAKNFNNISIDDAAMHPNDQPVWGNNGLOOQRYVAPVPO
POFNPSPMOESNDIEPNVNEALVGLSGSSSSQOMGIONTDYERKQIWGPPIISD
VOYPIOPHASFGEINSLGOSAPWSIGSOPAHFVWSNGMKRSDPFMKOQOPOHQ
YPMPLQGAWNTGMRHGPPEGQSGSPMMQODDTGGGGRNMGKRNMMNPMQOD
SNRMSGMPERNRPNHNOGRYSNGNGGVDPMPMDHMGKGNMGVAPNS
GHGSMKNNNAGGGGGGGGGRYNNRNGSGSGSGSNNMNSGNGSGSKMOF
MGSDQOYNNPEPIGGAADLTLYVMDHPNGELKKMOGRDITVYWGDELDPNDNNSK
IVGEAGDEDELTALNRCGPVPOKKGEDNORPPVPAKRPITVYWGDELDPNDNNSK
SESTIFDENNRNMDLTTEONPWLPHNOSTESSDWTSGMOCITPLNEPCTNHY
AVAEMLKNAVDKGTLDOSVTMMANLPVNLVYMLLVKIPALDSVENELKQITBESL
PEDVKEVDPQNOQKYNMDSOKLEHNRLLIEVYAKLEVOYSKRVNRLMADGIVPOQ
EQRAPASTEDYHYSFLE"
7929. .8084
/gene="C06G1.3"

CDS
7929. .8084
/gene="C06G1.3"
/note="coded for by C. elegans cDNA yk53b1.3; coded for by
C. elegans cDNA yk53b1.5"
/codon_start=1
/protein_id="AA82306.1"
/db_xref="GI:1086639"
/translation="MVFSRHCLLQTYRTSNKLIMGTVSOFETLSLPEICPPSLIS
HSSOSKS"
complement(8447. .8518)
/note="Asp: codon recognized: GAC"
complement(8851. .8922)
/note="Asp: codon recognized: GAC"
complement(9888. .9959)
/note="Asp: codon recognized: GAC"
12433. .13037
/gene="C06G1.2"
join(12433. .12799,12856. .13037)
/gene="C06G1.2"

CDS
/codon_start=1
/evidence=not_experimental
/protein_id="AA82307.1"
/db_xref="GI:1086640"
/translation="MNFALNSALFVILASASCOFMOYDVLGAVVSGVETEKAST
TEATPTTVVETTTVDSTTVEETPTTATTTATTTATTTATTTATTTATTTATTTAT
TTTAAATTTTTEPATTTTKEITTEITTEIVIAETTPTEVDEESLGFKKKDLATITRRL
FPRQHPGQKHVGOFGQOGF"
23882. .27748
/gene="C06G1.1"
join(23882. .23967,24973. .25153,25224. .25338,25698. .25769,
25834. .26091,26292. .26419,26517. .26642,26692. .26859,
27644. .27748)
/gene="C06G1.1"
/codon_start=1
/evidence=not_experimental
/protein_id="AA82308.1"
/db_xref="GI:1086641"
/translation="KCQIFRSDDLGVTCRTVEQCHATTPECKTNVELLMITGVNDSG

BASE COUNT	ORIGIN	CDS	trna	gene
9686 a	6087 t	5540 g	9892 t	

Query Match	79.0%	Score 15.8	DB 5	Length 31205
Best Local Similarity	89.5%	Pred. No. 7.8e+02		
Matches 17, Conservative	0	Mismatches 2	Indels 0	Gaps 0
Qy	2	aatgacgggaacgctccg	20	
db	24775	AATCACAGGAAACGTTCCG	24793	

RESULT	13
AC087901/c	
LOCUS	
DEFINITION	AC087901 . 47436 bp DNA HTG 06-FEB-2001 Mus musculus clone rp23-22123, WORKING DRAFT SEQUENCE, 17 unordered pieces.
ACCESSION	AC087901
VERSION	AC087901.1 GI:12667782
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 47436)
TITLE	Lau,C., Do,T. and Roe,B.A.
JOURNAL	Mus musculus BAC Clone rp23-22123
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 47436)
TITLE	Lau,C., Do,T. and Roe,B.A.
JOURNAL	Direct Submission
REFERENCE	Submitted (06-FEB-2001) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT	* NOTE: This is a 'working draft' sequence. It currently * consists of 17 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * 1 2423: contig of 2423 bp in length * 2424 2523: gap of unknown length. * 2524 4741: contig of 2218 bp in length * 4742 4841: gap of unknown length * 4842 6861: contig of 2020 bp in length

FEATURES	source
*	6862 6961: gap of unknown length
*	6962 9120: contig of 2159 bp in length
*	9121 9220: gap of unknown length
*	9221 11556: contig of 2336 bp in length
*	11557 11656: gap of unknown length
*	11657 14326: contig of 2670 bp in length
*	14327 14426: gap of unknown length
*	14427 16822: contig of 2396 bp in length
*	16823 16922: gap of unknown length
*	16923 19245: contig of 2323 bp in length
*	19246 19345: gap of unknown length
*	19346 21753: contig of 2408 bp in length
*	21754 21853: gap of unknown length
*	21854 25398: contig of 3545 bp in length
*	25399 25498: gap of unknown length
*	25499 26272: contig of 2774 bp in length
*	26273 28372: gap of unknown length
*	28373 30566: contig of 2194 bp in length
*	30567 30666: gap of unknown length
*	30667 33230: contig of 2564 bp in length
*	33231 33330: gap of unknown length
*	33331 35737: contig of 2407 bp in length
*	35738 35837: gap of unknown length
*	35838 39424: contig of 3587 bp in length
*	39425 39524: gap of unknown length
*	39525 43383: contig of 3859 bp in length
*	43384 43483: gap of unknown length
*	43484 47436: contig of 3953 bp in length.
Location/Qualifiers	
1..47436	
/organism="Mus musculus"	
/db_xref="taxon:10090"	
/clone="rp23-22123"	
/clone_lib="rp23"	

Query Match	79.04;	Score 15.8;	DB 78;	Length 47436;
Best Local Similarity	89.58;	Pred. NO.7.7e+02;		
Matches 17;	Conservative	0;	Mismatches 2;	Indels 0;
Gaps	0;			
1	caatgacggaaacgtccc	19		
Db	9739	CAATGACGGAAACGTTTC	9721	
RESULT 14				
AC021519/c				
LOCUS	AC021519	72593 bp	DNA	HTG
DEFINITION	Homo sapiens chromosome 11 clone RP11-319D2 map 11, LOW-PASS SEQUENCE SAMPLING.			
ACCESSION	AC021519			
VERSION	AC021519.2	GI:9148424		
KEYWORDS	HTG; HTGS_PHASE0.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	1 (bases 1 to 72593)			
TITLE	Bliren,B., Linton,L., Nusbaum,C. and Lander,E.			
JOURNAL	Homo sapiens chromosome 11, clone RP11-319D2			
REFERENCE	Unpublished			
AUTHORS	2 (bases 1 to 72593)			
	Bliren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Baldwin,J., Barra,N., Beckerly,R., Beda,F., Boguslavsky,L., Boukhalter,B., Brown,A., Burrell,G., Castle,A., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P., Darelanno,K., Dewar,K., Domino,M., Doyle,M., Fenster J., Ferreira,P., FlitzHugh,W., Forrest,C., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,T., Leloczky,J., Levine,R., Liu,C., Liu,G., Locke,K.,			

TITLE
JOURNAL

COMMENT

Macdonald, P., Marguis, N., McEwan, P., McGurk, A., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Olivar, T. M., Peterson, K.,
Pierre, N., Pisanic, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-JAN-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:6705581.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L5496

Center clone name: 319_D_2

NOTE: This record contains 75 individual

* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
844 943: gap of 100 bp in length
944 1807: contig of 864 bp in length
1808 1907: gap of 100 bp
1908 2770: contig of 863 bp in length
2771 2870: gap of 100 bp
2871 3770: contig of 900 bp in length
3771 3870: gap of 100 bp
3871 4722: contig of 852 bp in length
4723 4822: gap of 100 bp
4823 5702: contig of 880 bp in length
5703 5802: gap of 100 bp
5803 6667: contig of 865 bp in length
6668 6767: gap of 100 bp
6768 7657: contig of 890 bp in length
7658 7757: gap of 100 bp
7758 8637: contig of 880 bp in length
8638 8737: gap of 100 bp
8738 9612: contig of 875 bp in length
9613 9712: gap of 100 bp
9713 10583: contig of 871 bp in length
10584 10683: gap of 100 bp
10684 11555: contig of 872 bp in length
11556 11655: gap of 100 bp
11656 12512: contig of 857 bp in length
12513 12612: gap of 100 bp
12613 13467: contig of 875 bp in length
13468 13587: gap of 100 bp
13588 14468: contig of 881 bp in length
14469 14568: gap of 100 bp
14569 15428: contig of 860 bp in length
15429 15528: gap of 100 bp
15529 16398: contig of 870 bp in length
16399 16498: gap of 100 bp
16499 17379: contig of 881 bp in length
17380 17479: gap of 100 bp
17480 18364: contig of 885 bp in length
18365 18464: gap of 100 bp

18465 19354: contig of 890 bp in length
19355 19454: gap of 100 bp
19455 20340: contig of 886 bp in length
20341 20440: gap of 100 bp
20441 21310: contig of 870 bp in length
21311 21410: gap of 100 bp
21411 22303: contig of 893 bp in length
22304 22403: gap of 100 bp
22404 23256: contig of 853 bp in length
23257 23356: gap of 100 bp
23357 24243: contig of 887 bp in length
24244 24343: gap of 100 bp
24344 25189: contig of 846 bp in length
25190 25289: gap of 100 bp
25290 26151: contig of 862 bp in length
26152 26251: gap of 100 bp
26252 27073: contig of 822 bp in length
27074 27173: gap of 100 bp
27174 28052: contig of 879 bp in length
28053 28152: gap of 100 bp
28153 29054: contig of 902 bp in length
29055 29154: gap of 100 bp
29155 29964: contig of 810 bp in length
29965 30064: gap of 100 bp
30065 30946: contig of 882 bp in length
30947 31046: gap of 100 bp
31047 31926: contig of 880 bp in length
31927 32026: gap of 100 bp
32027 32901: contig of 875 bp in length
32902 33001: gap of 100 bp
33002 33840: contig of 839 bp in length
33841 33940: gap of 100 bp
33941 34821: contig of 881 bp in length
34822 34921: gap of 100 bp
34922 35795: contig of 874 bp in length
35796 35895: gap of 100 bp
35896 36723: contig of 828 bp in length
36724 36823: gap of 100 bp
36824 37709: contig of 886 bp in length
37710 37809: gap of 100 bp
37810 38681: contig of 872 bp in length
38682 38781: gap of 100 bp
38782 39644: contig of 863 bp in length
39645 39744: gap of 100 bp
39745 40630: contig of 886 bp in length
40631 40730: gap of 100 bp
40731 41580: contig of 850 bp in length
41581 41680: gap of 100 bp
41681 42555: contig of 875 bp in length
42556 42655: gap of 100 bp
42656 43471: contig of 816 bp in length
43472 43571: gap of 100 bp
43572 44433: contig of 862 bp in length
44434 44533: gap of 100 bp
44534 45401: contig of 868 bp in length
45402 45501: gap of 100 bp
45502 46365: contig of 864 bp in length
46366 46465: gap of 100 bp
46466 47335: contig of 870 bp in length
47336 47435: gap of 100 bp
47436 48286: contig of 851 bp in length
48287 48386: gap of 100 bp
48387 49278: contig of 892 bp in length
49279 49378: gap of 100 bp
49379 50265: contig of 887 bp in length
50266 50365: gap of 100 bp
50366 51235: contig of 870 bp in length
51236 51335: gap of 100 bp
51336 52218: contig of 883 bp in length
52219 52318: gap of 100 bp
52319 53193: contig of 875 bp in length
53194 53293: gap of 100 bp
53294 54147: contig of 854 bp in length


```
* 5418 54247: gap of 100 bp
* 5428 55137: contig of 890 bp in length
* 55138 55237: gap of 100 bp
* 55238 56111: contig of 874 bp in length
* 56112 56211: gap of 100 bp
* 56212 57071: contig of 860 bp in length
* 57072 57171: gap of 100 bp
* 57172 58030: contig of 859 bp in length
* 58031 58130: gap of 100 bp
* 58131 59013: contig of 883 bp in length
* 59014 59113: gap of 100 bp
* 59114 59982: contig of 869 bp in length
* 59983 60082: gap of 100 bp
* 60083 60966: contig of 884 bp in length
* 60967 61066: gap of 100 bp
* 61067 61935: contig of 869 bp in length
* 61936 62035: gap of 100 bp
* 62036 62909: contig of 874 bp in length
* 62910 63009: gap of 100 bp
* 63010 63816: contig of 807 bp in length
* 63817 63916: gap of 100 bp
* 63917 64804: contig of 888 bp in length
* 64805 64904: gap of 100 bp
* 64905 65793: contig of 881 bp in length
* 65796 65895: gap of 100 bp
* 65896 66766: contig of 871 bp in length
* 66767 66866: gap of 100 bp
* 66867 67702: contig of 836 bp in length
* 67703 67802: gap of 100 bp
* 67803 68676: contig of 874 bp in length
```

Query Match 79.0%; Score 15.8; DB 67; Length 72593;
Best Local Similarity 89.5%; Pred. No. 7.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 caatgacgggaaagcttc 19
|||||
Db 66710 CAATGTCGGGAATCGTCC 66692

RESULT 15
LMFLCHR34_1/c
WPCOMMENT

Sequence split into 7 fragments LOCUS LMFLCHR34 Accession AL499623

Fragment Name	Begin	End
LMFLCHR34_0	1	110000
LMFLCHR34_1	100001	210000
LMFLCHR34_2	200001	310000
LMFLCHR34_3	300001	410000
LMFLCHR34_4	400001	510000
LMFLCHR34_5	500001	610000
LMFLCHR34_6	600001	697361

Continuation (2 of 7) of LMFLCHR34 from base 100001 (AL499623 *Leishmania* major chromosome)

Query Match 79.0%; Score 15.8; DB 84; Length 110000;
Best Local Similarity 89.5%; Pred. No. 7.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 caatgacgggaaagcttc 19
|||||
Db 46221 CAACGACGGGGAACGTCC 46203

Search completed: June 8, 2001, 06:03:00
Job time: 7310 sec

• • • • •

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 8, 2001, 06:21:25 ; Search time 217.33 seconds
(without alignments)
53.723 Million cell updates/sec

Title: US-09-507-242-6

Perfect score: 20

Sequence: 1 caatgacgggaacgttcg 20

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 678276 seqs, 291890651 residues 1356552

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N.Geneseq_0401.*

1: /SID56/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	A50294	Candida albicans C
2	20	100.0	989	A50292	Candida albicans C
3	15.2	76.0	274	T24451	Human gene signatu
4	15.2	76.0	479	Z91619	E.coli promoter re
5	15.2	76.0	560	C04556	Human secreted pro
6	15.2	76.0	603	C42788	Arabidopsis thalia
7	15.2	76.0	1062	C48389	Arabidopsis thalia
8	15.2	76.0	1315	F13745	Aspergillus oryzae
9	15.2	76.0	1473	X61690	B. burgdorferi ant
10	15.2	76.0	1527	X61689	B. burgdorferi ant
11	15.2	76.0	3107	C77447	Human ORFX ORF3002

12	15.2	76.0	7801	13	Q23500	Zymomonas xylose 1
13	15	75.0	1589	20	X84935	Human secreted pro
14	14.8	74.0	1321	20	X25655	Human endogenous r
15	14.8	74.0	1321	21	A59205	5' non coding and
16	14.8	74.0	1422	21	C38943	Arabidopsis thalia
17	14.8	74.0	1539	21	C42085	Arabidopsis thalia
18	14.8	74.0	2372	20	X25664	Human endogenous r
19	14.8	74.0	2372	21	A59214	R-U5 region and pa
20	14.8	74.0	2422	20	Z00844	Human secreted pro
21	14.8	74.0	2938	20	X25656	Human endogenous r
22	14.8	74.0	2938	21	A59206	Gag and partial po
23	14.8	74.0	3531	21	Z36787	DNA sequence of th
24	14.8	74.0	6723	20	X13100	Enterococcus faeca
25	14.8	74.0	7582	20	X25665	Complete human end
26	14.8	74.0	7582	21	A59215	Human endogenous r
27	14.8	74.0	10636	20	X13011	Enterococcus faeca
28	14.8	74.0	138169	21	A34791	Human adenosine re
29	14.8	74.0	141589	21	F20913	Human ELAM-1 polyn
30	14.8	74.0	141589	21	F21127	Human low adenosin
31	14.8	74.0	141589	21	F21152	Human low adenosin
32	14.8	74.0	141589	21	A35005	Human adenosine re
33	14.8	74.0	141589	21	A35030	Human adenosine re
34	14.8	74.0	146981	21	F21442	Human ELAM-1 polyn
35	14.8	74.0	209273	21	F21437	Human factor-relat
36	14.6	73.0	818	13	Q32502	HCY NS4-NS5 clone
37	14.4	72.0	619	21	F11362	Aspergillus niger
38	14.2	71.0	334	21	C34374	Arabidopsis thalia
39	14.2	71.0	467	21	A30458	Human brain cDNA c
40	14.2	71.0	603	21	A10742	Fragment #1 of C.
41	14.2	71.0	743	21	F13533	Aspergillus oryzae
42	14.2	71.0	797	21	F08035	Fusarium venenatum
43	14.2	71.0	873	20	Z16753	Human gene express
44	14.2	71.0	882	19	V40176	Flavobacterium sp.
45	14.2	71.0	1157	21	C34054	Arabidopsis thalia

ALIGNMENTS

RESULT 1	A50294	standard; DNA; 20 BP.
ID	A50294	
AC	A50294	
XX	20-NOV-2000	(first entry)
DE	Candida albicans Caess1 gene-specific primer OM-221.	
XX	Caess1; ESS1; Infection; diagnosis; therapy; antifungal; fungicide;	
KW	antiproliferative; cytostatic; antitumour; PCR primer; ss.	
OS	Candida albicans.	
XX	WO200050561-A2.	
PN	31-AUG-2000.	
PD	18-FEB-2000; 2000WO-US04203.	
PF	23-FEB-1999; 9905-0121246.	
PR	(HEAL-) HEALTH RES INC.	
PA	Hanes SD, Devasahayam G, Chaturvedi V;	
XX	WPI, 2000-565453/52.	
DR	Novel Candida albicans gene, Caess1 useful for identifying compounds	
PT	that specifically bind to and/or inhibit Caess1 and thus for treating	
PT	Candida albicans infections and other life-threatening fungal	
PT	Infections	
XX		

PS Claim 6; Page 38; 51pp; English.

XX This is the nucleotide sequence of PCR primer OW-221, which is
CC specific to the CaESS1 gene (see A50292), the Candida albicans
CC homologue of *Saccharomyces cerevisiae* ESS1. OW-221 can be used
CC with primer OW-216 (see A50293) for the specific amplification of
CC CaESS1 DNA; they do not amplify ESS1 DNA or PIN1 (the human
CC homologue of ESS1) DNA. Detection of CaESS1 nucleic acids can
CC be used as a means for diagnosing *C. albicans* infection.

XX Sequence 20 BP; 6 A; 5 C; 6 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.076;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatgacgaggaacgtccg 20

Db 1 caatgacgaggaacgtccg 20

RESULT 2

A50292/C
ID A50292 standard; DNA; 989 BP.

XX A50292;

DT 20-NOV-2000 (first entry)

XX Candida albicans CaESS1 gene.

XX CaESS1; ESS1; infection; diagnosis; therapy; antifungal; fungicide;

KW antiproliferative; cytostatic; antitumour; ds.

XX Candida albicans.

XX Location/Qualifiers

Key 256..789

FT CDS

XX MO200050561-A2.

XX 31-AUG-2000.

PE 18-FEB-2000; 2000WO-US04203.

XX 23-FEB-1999; 99US-0121246.

XX (HEAL-) HEALTH RES INC.

PI Hanes SD, Devasahayam G, Chaturvedi V;

DR WPI: 2000-565453/52.

DR P-PSDB: Y95876.

PT Novel Candida albicans gene, CaESS1 useful for identifying compounds
PT that specifically bind to and/or inhibit CaESS1 and thus for treating
PT Candida albicans infections and other life-threatening fungal
PT infections -

PS Claim 2; Fig 1A; 51pp; English.

XX This is the complete nucleotide sequence of the CaESS1 gene from
CC Candida albicans. The gene encodes a 177-amino acid protein,
CC CaESS1 (see Y95876), which is the *C. albicans* homologue of
CC *Saccharomyces cerevisiae* ESS1 (428 amino acid identity). The
CC CaESS1 gene was isolated from a *C. albicans* genomic DNA library by
CC functional complementation of a temperature-sensitive *S. cerevisiae*
CC strain, *ess1-194pts*. CaESS1 nucleic acids, especially CaESS1-specific
CC primers (see A50293-94) and probes to determine the presence of *C.*
CC albicans in a sample or specimen. CaESS1 protein is a target
CC for screening for antifungal and inhibitor compounds, useful for

CC treating or preventing *C. albicans* infections. CaESS1 DNA can also
CC be used to generate diagnostic probes or primers for replicating or
CC cloning *C. albicans* DNA. A CaESS1 inhibitor can be a compound
CC which selectively inhibits growth of *S. cerevisiae* not containing an
CC endogenous ESS1 gene but rather CaESS1 and uninduced PIN1 (a human
CC homologue of ESS1) and/or preferably does not inhibit induced PIN1,
CC e.g. does not inhibit *S. cerevisiae* not containing an endogenous
CC ESS1 but rather induced PIN1. Compositions which inhibit PIN1
CC are useful antiproliferatives e.g. antineoplastics, antitumour
CC agents or anticancer agents.

XX Sequence 989 BP; 340 A; 174 C; 190 G; 285 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 989;

Best Local Similarity 100.0%; Pred. No. 0.12;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 caatgacgaggaacgtccg 20

Db 970 CAATGACGAGGAACGTCCG 951

RESULT 3

T24451/C
ID T24451 standard; cDNA to mRNA; 274 BP.

XX T24451;

DT 10-OCT-1996 (first entry)

XX Human gene signature HUMGS06487.

XX Gene signature; messenger RNA; mRNA; relative abundance; frequency;

KW human; cloning; mapping; non-biased library; diagnosis; detection;

KW cell typing; abnormal cell function; ss.

XX Homo sapiens.

XX MO9514772-A1.

XX 01-JUN-1995.

PE 11-NOV-1994; 94WO-JP01916.

XX 12-NOV-1993; 93JP-0355504.

XX (MATS/) MATSUBARA K.

PA (OKUB/) OKUBO K.

PI Matsubara K, Okubo K;

DR WPI: 1995-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 1613; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be

XX C42788;
AC
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 36852.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 603;
Best Local Similarity 85.0%; Pred. No. 46;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 caatgacggaaaccttcg 20
||||| ||||| ||||| |||||
Db 274 CAATGACGAGAAACGTGCTG 255

RESULT 7
ID C48389 standard; DNA; 1062 BP.
XX C48389;
XX 18-OCT-2000 (first entry).

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 57303.
DE

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PD 06-SEP-2000.
XX

PE 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135359.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.

PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158322.
PR 11-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161044.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.0%; Score 15.2; DB 21; Length 1062;
Best Local Similarity 85.0%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 caatgacggaacgtccg 20
Db 733 CAATGACGAGAACGTCTG 714

RESULT 8
F13745/C
ID F13745 standard; CDNA; 1315 BP.

XX F13745;
AC
XX
XX 13-MAR-2001 (first entry)
XX
XX
DE Aspergillus oryzae EST SEQ ID NO:6268.

KW Multiple gene expression; filamentous fungal cell; EST;
KW expressed sequence tag; Fusarium venenatum; Aspergillus niger;
KW Aspergillus Oryzae; Trichoderma reesei; identification; recombination;
KW culture condition; environmental stress; spore morphogenesis;
KW metabolic pathway engineering; catabolic pathway engineering; ss.
XX
OS Aspergillus oryzae.


```

XX XX WO200056762-A2.
XX PN
XX XX
XX XX 28-SEP-2000.
XX PD
XX XX
XX XX 22-MAR-2000; 2000WO-US07781.
XX PF
XX XX 22-MAR-1999; 99US-0273623.
XX PR
XX XX (NOVO ) NOVO NORDISK BIOTECH INC.
XX PA (NOVO ) NOVO NORDISK AS.
XX PA
XX PI Berka RM, Rey MW, Shuster JR, Kauplinen S, Clausen IG, Olsen PB;
XX XX
XX XX WPI: 2000-594572/56.
XX DR
XX XX
XX XX Monitoring differential expression of genes in filamentous fungal cells
XX PT uses fluorescence-labeled nucleic acids isolated from the cells and a
XX PT substrate of expressed sequence tags -
XX PS
XX XX Claim 88: Page 2577-2578; 3161pp; English.
XX CC
XX CC The present invention describes a method for monitoring differential
XX CC expression of genes in a first filamentous fungal (FF) cell relative to
XX CC expression of the same genes in one or more second filamentous fungal
XX CC cells. The method uses fluorescence-labeled nucleic acids isolated from
XX CC the FF cells and a substrate of expressed sequence tags (EST). The ESTs
XX CC are used in the methods for monitoring differential expression of genes
XX CC in a first filamentous fungal (FF) cell relative to expression of the
XX CC same genes in one or more second filamentous fungal cells. Monitoring
XX CC the global expression of genes from FF cells allows the production
XX CC potential of the microorganisms to be improved. New genes may be
XX CC discovered, possible functions of unknown open reading frames can be
XX CC identified and gene copy number variation and stability can be
XX CC monitored. The expression of genes can be used to study how FF cells
XX CC adapt to changes in culture conditions, environmental stress, spore
XX CC morphogenesis, recombination, metabolic or catabolic pathway engineering.
XX CC Using ESTs provides several advantages over genomic or random cDNA
XX CC clones including elimination of redundancy as one spot on an array
XX CC equals one gene or open reading frame, and organisation of the
XX CC microarrays based on function of the gene products to facilitate
XX CC analysis of the results. F07478 to F1247 represents ESTs from Fusarium
XX CC venenatum; F11248 to F11853 represents ESTs from Aspergillus niger;
XX CC F11854 to F14878 represents ESTs from Aspergillus oryzae; and F14879 to
XX CC F15337 represents ESTs from Trichoderma reesei, which are all
XX CC specifically claimed in the present invention.
XX CC
XX XX Sequence 1315 BP; 305 A; 370 C; 262 G; 378 T; 0 other;
XX SQ

```

Query Match 76.0%; Score 15.2; DB 21; Length 1315;
 Best Local Similarity 85.0%; Pred. No. 50;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 caatgacgggaaagcttcgcg 20
   ||| ||| ||| ||| ||| |||
Db 954 CAAGACACGGAACGCTCCG 935

```

RESULT 9
 X61690
 ID X61690 standard; DNA; 1473 BP.
 XX
 XX AC X61690;
 XX
 XX DT 19-JUL-1999 (first entry)
 XX DT
 XX DE B. burgdorferi antigenic protein coding sequence, t752.nt.
 XX XX
 XX KW Antigenic protein; vaccine; Lyme disease; Infection; detection; ss.
 XX XX
 XX OS Borrelia burgdorferi.
 XX XX

```

PN PM WO9859071-A1.
XX XX
XX XX 30-DEC-1998.
XX PD
XX XX
XX XX 18-JUN-1998; 98WO-US12718.
XX PF
XX XX 03-SEP-1997; 97US-0057483.
XX PR 20-JUN-1997; 97US-0050359.
XX PR 22-JUL-1997; 97US-0053344.
XX PR 22-JUL-1997; 97US-0053377.
XX XX
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PA (MEDI-) MEDIMUNE INC.
XX PI Choi GH, Erwin AL, Hanson MS, Lathigra R;
XX XX
XX XX WPI: 1999-189980/16.
XX DR
XX XX P-PSDB: Y19993.
XX PT
XX PT New isolated Borrelia burgdorferi nucleic acids - used to develop
XX PT products for the diagnosis, prevention and treatment of diseases
XX PT caused by Borrelia, particularly Lyme disease
XX PS
XX XX Claim 1: Page 154; 275pp; English.
XX CC
XX CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
XX CC invention, which is suitable for use in a vaccine. The Bb polypeptides
XX CC can be used in vaccines for eliciting protective antibodies to members of
XX CC the Borrelia genus, particularly for the use against Lyme disease in
XX CC humans and animals. They can be used for preventing or attenuating an
XX CC infection caused by a member of the Borrelia genus. The products can also
XX CC be used for detection of members of the Borrelia genus.
XX XX
XX SQ Sequence 1473 BP; 468 A; 166 C; 290 G; 549 T; 0 other;

```

Query Match 76.0%; Score 15.2; DB 20; Length 1473;
 Best Local Similarity 85.0%; Pred. No. 51;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

OY 1 caatgacgggaaagcttcgcg 20
   ||| ||| ||| ||| ||| |||
Db 300 caatgacgggaaagcttcgcg 319

```

RESULT 10
 X61689
 ID X61689 standard; DNA; 1527 BP.
 XX
 XX AC X61689;
 XX
 XX DT 19-JUL-1999 (first entry)
 XX DT
 XX DE B. burgdorferi antigenic protein coding sequence, f752.nt.
 XX XX
 XX KW Antigenic protein; vaccine; Lyme disease; Infection; detection; ss.
 XX XX
 XX OS Borrelia burgdorferi.
 XX XX
 XX PN WO9859071-A1.
 XX PD
 XX PD 30-DEC-1998.
 XX PF
 XX PF 18-JUN-1998; 98WO-US12718.
 XX PR
 XX PR 03-SEP-1997; 97US-0057483.
 XX PR 20-JUN-1997; 97US-0050359.
 XX PR 22-JUL-1997; 97US-0053344.
 XX PR 22-JUL-1997; 97US-0053377.
 XX XX
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA (MEDI-) MEDIMUNE INC.
 XX XX

PI Choi GH, Erwin AL, Hanson MS, Iathigra R;
 XX MPI: 1999-189980/16.
 DR P-PSDB; Y19992.
 XX
 PT New isolated Borrelia burgdorferi nucleic acids - used to develop
 PT products for the diagnosis, prevention and treatment of diseases
 PT caused by Borrelia, particularly Lyme disease
 PS
 XX Claim 1: Page 153-154; 275pp; English.
 CC This sequence encodes a Borrelia burgdorferi (Bb) protein of the
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides
 CC can be used in vaccines for eliciting protective antibodies to members of
 CC the Borrelia genus, particularly for the use against Lyme disease in
 CC humans and animals. They can be used for preventing or attenuating an
 CC infection caused by a member of the Borrelia genus. The products can also
 CC be used for detection of members of the Borrelia genus.
 CC
 SQ Sequence 1527 BP; 496 A; 168 C; 296 G; 567 T; 0 other;
 Query Match 76.0%; Score 15.2; DB 20; Length 1527;
 Best Local Similarity 85.0%; Pred. No. 51;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 caatgacggaagaacttcgcg 20
 Db 354 cactgacggaagaacttcgtg 373
 RESULT 11
 C77447/c
 ID C77447 standard; cDNA; 3107 BP.
 AC C77447;
 XX
 DT 08-FEB-2001 (first entry)
 DE Human ORFX ORF3002 polynucleotide sequence SEQ ID NO:6003.
 XX
 KW Human: open reading frame: ORFX; detection: cytostatic; hepatotropic;
 KW vulnerrary; antiprosoriatric; antiparkinsonian; neurotropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antihypoid;
 KW antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO200058473-A2.
 PD 05-OCT-2000.
 XX
 PF 31-MAR-2000; 2000MO-US08621.
 XX
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinkets RA, Leach M;

DR MPI: 2000-602362/57.
 DR P-PSDB; B43238.
 XX
 PT Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 PS
 XX Claim 5: Page 5182-5184; 5507pp; English.
 CC C74446 to C77606 encode the proteins given in B40237 to B43397, which
 CC represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytosstatic; hepatotropic; vulnerrary;
 CC antiprosoriatric; antiparkinsonian; neurotropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antihypoid; and antianaemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, and antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive.
 CC
 SQ Sequence 3107 BP; 546 A; 965 C; 981 G; 613 T; 2 other;
 Query Match 76.0%; Score 15.2; DB 21; Length 3107;
 Best Local Similarity 85.0%; Pred. No. 55;
 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 OY 1 caatgacggaagaacttcgcg 20
 Db 2459 CACTGACGGAAGAACTTCGC 2440
 RESULT 12
 Q23300
 ID Q23500 standard; DNA; 7801 BP.
 AC Q23500;
 XX
 DT 26-AUG-1992 (first entry)
 DE Zymomonas xylose isomerase expression plasmid pX1PX.
 XX
 KW xylA; xylB; pZM1A; pyruvate decarboxylase; promoter; marker gene;
 KW chloramphenicol resistance; ss.
 XX
 OS Zymomonas mobilis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1618..2940
 FT /*tag- a
 FT /product= xylose_isomerase
 FT /note= "see R22720"
 FT 3013..4468
 FT /*tag- b
 FT /standard_name= xylB
 FT /note= "xylokinase"
 FT 1521..1592
 FT /*tag- c
 FT /label= pyruvate_decarboxylase_promoter
 FT 1531..1543
 FT /*tag- d
 FT 1559..1568
 FT /*tag- e
 FT -10_signal
 FT -35_signal
 FT -10_signal

```
FT misc_feature 1..219
FT /tag= f
FT /phenotype= chloramphenicol_resistance
FT 7361..7801
FT **tag= g
FT /phenotype= chloramphenicol_resistance
FT 5414..7096
FT /tag= h
FT /label= pm21A
FT /note= "when included in a plasmid this region
FT allows replication in Zymomonas"
FT
FT JP04066090-A.
FT
FT 02-MAR-1992.
FT
FT 05-JUL-1990; 90JP-0176420.
FT
FT 05-JUL-1990; 90JP-0176420.
FT
FT (ENER-) SHIN ENERGY SANGYO.
FT (NENR-) NENRYOYO ALCOHOL KAIHAT.
FT
FT WPI: 1992-120685/15.
FT P-PSDB: R22720.
FT
FT Vector for expression of heterogenes in Zymomonas sp.
FT used for expression of xylose isomerase for prodn. of ethanol from
FT xylose and/or fructose
FT
FT Disclosure: Fig 1: 14pp; Japanese.
FT
FT This is the sequence of plasmid pX1PX. The plasmid can be used to
FT transform Zymomonas cells. The transformants can be cultured to
FT express the xylA/B genes. Alcohol production from such
FT transformants is comparable to that from yeast.
FT
FT Sequence 7801 BP; 1959 A; 1910 C; 1918 G; 2014 T; 0 other:
SQ
Query Match 76.0%; Score 15.2; DB 13; Length 7801;
Best Local Similarity 85.0%; Pred. No. 62;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 caatgacgggaacgttcg 20
DB 4410 ccaagcagcgaagaacgttcg 4429
RESULT 13
X84935/c
ID X84935 standard; DNA: 1589 BP.
XX
AC X84935:
XX
DT 30-JUL-1999 (first entry)
XX
DE Human secreted protein gene No. 3.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN MO9924836-A1.
XX
PD 20-MAY-1999.
```

```
XX
XX 04-NOV-1998; 98WO-US23435.
XX
XX 17-NOV-1997; 97US-0066100.
XX
XX 07-NOV-1997; 97US-0064900.
XX
XX 07-NOV-1997; 97US-0064908.
XX
XX 07-NOV-1997; 97US-0064911.
XX
XX 07-NOV-1997; 97US-0064912.
XX
XX 07-NOV-1997; 97US-0064983.
XX
XX 07-NOV-1997; 97US-0064984.
XX
XX 07-NOV-1997; 97US-0064985.
XX
XX 07-NOV-1997; 97US-0064987.
XX
XX 07-NOV-1997; 97US-0064988.
XX
XX 17-NOV-1997; 97US-0066090.
XX
XX 17-NOV-1997; 97US-0066094.
XX
XX 17-NOV-1997; 97US-0066095.
XX
XX 17-NOV-1997; 97US-0066089.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Carter KC, Ebdner R, Endress GA, Peng P, Janat F;
XX Ryan H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
XX Ruben SM, Shi Y, Soppet DR, Wei Y;
XX
XX WPI: 1999-337740/28.
XX P-PSDB: Y27569.
XX
XX New human secreted proteins and coding sequences useful for treating
XX disorders of the immune system and hyperproliferative disorders
XX
XX Claim 1: Page 283; 507pp; English.
XX
XX This sequence represents a nucleic acid molecule which encodes a secreted
XX human protein. The gene number is given in the descriptor line. The gene
XX can be used to generate fusion proteins by linking to the gene to a human
XX immunoglobulin Fc portion (e.g. X84924) for increasing the stability of
XX the fused protein as compared to the human protein only.
XX
XX The invention relates to 125 novel genes and their fragments (nucleic
XX acid sequences: X84933-X85057; amino acid sequences Y27567-Y27933) which
XX are useful for preventing, treating or ameliorating medical conditions
XX e.g. by protein or gene therapy. Also, pathological conditions can be
XX diagnosed by determining the amount of the new polypeptides in a sample
XX or by determining the presence of mutations in the new polynucleotides.
XX Specific uses are described for each of the 125 polynucleotides, based on
XX which tissues they are most highly expressed in (see X84933 for described
XX uses).
XX
XX Sequence 1589 BP; 359 A; 379 C; 361 G; 478 T; 12 other:
SQ
Query Match 75.0%; Score 15; DB 20; Length 1589;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 6 acggggaacgttcg 20
DB 355 ACGGGAACGTTCCG 341
RESULT 14
X25655
ID X25655 standard; cDNA to mRNA: 1321 BP.
XX
AC X25655:
XX
DT 21-MAY-1999 (first entry)
XX
DE Human endogenous retrovirus W clone cl.6A2.
XX
KW Clone; human endogenous retrovirus; genome; autoimmune disease;
KW multiple sclerosis; rheumatoid polyarthritis; insulin-dependent diabetes;
KW disseminated lupus erythematosus; pregnancy; chromosomal marker; ss.
XX
```


**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☒ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (usp...)